TABLE OF CONTENTS

PAGE

MAY 2 1980

NASA Contractor Report 3257

A Model-Based Technique for Predicting Pilot Opinion Ratings for Large Commercial Transports

William H. Levison

NASA

(79)

NASA Contractor Report 3257

A Model-Based Technique for Predicting Pilot Opinion Ratings for Large Commercial Transports

William H. Levison
Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

NASA

National Aeronautics
and Space Administration

Scientific and Technical
Information Office

1980

BLANK PAGE

ii

TABLE OF CONTENTS

LIST OF TABLES

PAGE

LIST OF FIGURES

SUMMARY

A model-based technique for predicting pilot opinion ratings is described. Features of this procedure, which is based on the optimal-control model for pilot/vehicle systems, include (1) capability to treat "unconventional" aircraft dynamics, (2) a relatively free-form pilot model, (3) a simple scalar metric for attentional workload, and (4) a straightforward manner of proceeding from descriptions of the flight task environment and requirements to a prediction of pilot opinion rating. The method was able to provide a good match to a set of pilot opinion ratings obtained in a manned simulation study of large commercial aircraft in landing approach.

1

1. INTRODUCTION

Manufacturers of commercial aircraft require more general and more reliable methods of predicting aircraft handling qualities than currently exist. Existing criteria have been developed primarily for military aircraft and have been validated largely for high-performance aircraft such as fighters. At present, reliable techniques for extending existing criteria to large commercial transports are not available.

This report presents the results of a study performed by Bolt Beranek and Newman Inc. (BBN), with the aid of Douglas Aircraft Company (Douglas), to develop and test a model-based technique for predicting the influence of aircraft response parameters and other relevant factors on pilot opinion ratings. While the procedure is intended to have general application, the focus in this report is on large transports. Frequent reference is made to a manned simulation study performed by Douglas in 1975.* To facilitate discussion, the analytic study that is the subject of this report will be referred to as the "BBN study", whereas the preceding simulation program will be referred to as the "Douglas study".

1.1 Vehicle-Centered Handling Qualities Criteria

The determination of quantitative requirements for handling qualities is a complex task. Handling characteristics must be specified in the operational, service, and permissible flight envelopes. Situations that can only be described in a statistical sense (e.g., failures, turbulence) must also be considered. The task is especially difficult, because objective requirements must be determined that correlate well with subjective pilot assessments.

---

* This effort included a subcontract to Douglas Aircraft Company to provide a data base extracted from the 1975 Douglas simulation study and to provide other consulting services. Mr. William W. Rickard was project engineer for the Douglas effort.

Open-loop vehicle behavior can be determined in a relatively straightforward manner from in-flight tests. For this reason, handling qualities specifications are based almost exclusively on open-loop vehicle response characteristics [1]. Criteria are specified for both transient response and frequency response characteristics.

Predictions of handling qualities based on frequency-domain characteristics are generally obtained from the denominator characteristics of the linearized vehicle equations of motion. Natural frequencies and damping ratios of dominant pole-pairs, time constants of real poles, and, in some cases, numerator time constants, are first estimated. These values are then checked against tables and charts which purport to show the relationship between handling quality "levels" and parameter values considered in combination as well as singly.

If handling qualities requirements are based on vehicle denominator characteristics alone, the aircraft manufacturer can evaluate the performance of his aircraft in this regard through a series of relatively straightforward in-flight tests. He need not be concerned with the interaction between the vehicle and the pilot, which, of course, will vary from one test pilot to the next. The ease with which compliance can be tested, plus the existence of a substantial body of relevant handling qualities data, provide a strong motivation to relate handling qualities to open-loop frequency-response characteristics.

Nevertheless, despite its relative convenience, this procedure has some very serious limitations:

   a. the procedure is lacking in generality, since it requires that an extensive set of tables and charts be available. The data obtained in these charts are applicable only to vehicle dynamics that closely resemble the type of aircraft used in constructing the data base. Accordingly, techniques based on

frequency-domain analysis of open-loop vehicle response cannot be applied with confidence to systems where new force producers and stability augmentation systems may significantly modify the effective vehicle dynamics.

b. In general, the numerator characteristics of the vehicle transfer functions are given considerably less attention than the denominator characteristics. Nevertheless, numerator characteristics may have an important influence on handling qualities.

The open-loop response of the vehicle to a step control input is often used to assess vehicle handling qualities. One particular response pattern that is the so-called C^* response, which is a combination of normal acceleration, pitch rate, and pitch acceleration responses to a step input of stick force. Boundaries on this response waveform are associated with a pilot rating of 3.5. That is, if the response of the actual or simulated vehicle exceeds this envelope at any point, the vehicle should be rated as less than "satisfactory" on the Cooper rating scale.

This technique is somewhat more general than the frequency-domain procedure discussed above in that the effects of zeros as well as poles in the vehicle transfer function are considered. Nevertheless, the generality of this method is somewhat in doubt, since unreliable predictions have been reported in studies of "non-standard" vehicle dynamics [2]. Furthermore, the C^* analysis procedure yields a binary result - the predicted Cooper rating is either higher or lower than 3.5. Consequently, this scheme does not allow one to estimate the extent of the degradation in handling qualities when the vehicle response is outside the prescribed envelope for only a short time.

Frequency-domain and time-domain procedures confined to open-loop vehicle analysis exhibit the following additional deficiencies:

4

a. Effects of atmospheric disturbances such as turbulence and wind shears are not considered.* It has been shown, nevertheless, that pilot opinion can be influenced appreciably by the action of disturbances external to the vehicle [3].

b. Effects of displays (such as flight directors) are not considered. To the extent that display parameters influence overall mission suitability (and, therefore, pilot opinion rating), a method for predicting handling qualities should account for the effects of display parameters.

Application of existing techniques for predicting handling qualities of large commercial transports is limited in a number of ways. For example:

a. Existing handling qualities criteria have been developed primarily for military aircraft. Furthermore, these criteria have been validated largely for high-performance aircraft (fighters, etc.). Thus, application to large commercial transports cannot be undertaken with great confidence.

b. Most existing criteria are based on simple models of aircraft dynamics in which phugoid and short-period response characteristics can be distinguished. Consequently, application to aircraft having relaxed static stability and substantial control augmentation is dubious at best.

c. For the most part, effects of turbulence are not considered. This oversight neglects a potentially important aspect of flying qualities and is a consequence of considering only open-loop aircraft characteristics.

---

* The handling qualities specifications of [1] do impose a few requirements with regard to atmospheric disturbances, and analytic models for these disturbances are specified. Suggestions have been offered for updating this section of the military specification [4].

d. Present methods do not consider effects of dynamic aeroelasticity.

1.2 Model-Based Schemes for Predicting Handling Qualities

Pilot/vehicle analysis can allow considerably greater insight into the handling qualities of an aircraft control system than can be obtained by analysis of open-loop vehicle response alone. Predictions can be obtained for closed-loop response (which is usually what counts in terms of meeting mission requirements), and the demands made on the pilot can be explored. Thus, the effects of external disturbances and control/display parameters, as well as inherent pilot limitations, can be considered.

Reasonably general models for pilot response characteristics are available. That is, by specifying a relatively small number of constraints, one can predict pilot response characteristics for a wide variety of control situations. Consequently, predictive schemes based on pilot/vehicle analysis are not constrained to deal with "conventional" dynamics and are thus potentially more general than techniques based solely on open-loop vehicle characteristics.

Until recently, application of pilot/vehicle analysis to studies of vehicle handling qualities has been based primarily on servo-theory techniques. Central to these techniques is a frequency-domain model of the pilot which is generally structured so that feedback loops are closed serially, rather than in parallel. Typically, the pilot's control strategy for each loop is represented by a low-frequency gain, a lead-lag network, and an equivalent time delay to represent inherent information-processing delays. (Usually, pilot neuromuscular lags are neglected or are incorporated into the effective time delay.)

Analysis of the pilot/vehicle system is based on the assumption that the pilot attempts to achieve "good" performance in terms of the gain-crossover frequency and phase margin associated with each control loop. Ideally, crossover frequencies

are kept sufficiently large to assure adequate response bandwidth while comfortably large phase margins and damping ratios are maintained in order to assure high-frequency stability. By use of root locus techniques, a set of pilot gains and lead time constants is found which best satisfies these requirements. If the closed-loop frequency response is not within the desired envelope, or if substantial pilot lead generation is required, then the pilot rating has to be degraded to take these factors into account [2].

Perhaps the most comprehensive effort to apply classical control theory to the prediction of aircraft handling qualities has been conducted by R. O. Anderson and his associates in the development of the "Paper Pilot" analysis scheme [5]. This scheme relates pilot rating to metrics of both closed-loop system performance and pilot workload, and it introduces the concept that the pilot operates so as to minimize his rating score.

Pilot rating is assumed to be an explicit function of system performance and pilot lead requirements as given by the following expression:

$$R = 1 + K_\sigma \left[\frac{\sigma - \sigma_o}{\sigma} \right] + \Sigma_i \, K_{L_i} \cdot T_{L_i}$$

(1)

where R is the predicted Cooper rating, σ is some measure of overall system performance (say, a linear combination of rms variations in flight path and attitude variables), T_{L_i} is the lead time constant generated by the pilot in the ith control loop, and K_σ and K_{L_i} are weighting coefficients. The variable σ_o represents the desired performance level in a particular task. System performance degrades the rating only when $\sigma > \sigma_o$.

A pilot model of the form described above is used in this scheme, and pilot parameters are found which minimize the

predicted rating. Since there is generally a trade-off between lead generation and system performance, the predicted pilot behavior depends on the choice of coefficients in Equation (1). If a relatively high weighting is placed on minimizing system errors, large lead time constants are likely to result. Conversely, by increasing the penalty associated with lead equalization, lower lead time constants and increased system errors will be predicted. Good matches to experimental data have been obtained for a variety of control tasks through appropriate formulation of the rating expression and adjustment of the weighting coefficients [5-8].

This analysis scheme allows one to account for some of the factors (other than open-loop vehicle response characteristics) that influence pilot opinion. Pilot compensation and gain requirements are determined directly, and the susceptibility of the system to PIO's can be estimated from the closed-loop pole-zero and Bode plots. Effects of external disturbances, and to some extent display parameters, are accounted for.

Perhaps the most serious limitation of the Paper Pilot scheme is that no general rule has yet been determined for choosing the precise form of the rating expression or for selecting the various weighting coefficients. Other limitations inherent in servo-analysis techniques are also important:

a. A relatively constrained fixed-form pilot model is usually employed; the structure of this model and nature of loop closures must be assumed a priori to the analysis. This procedure is not straightforward when applied to multi-loop control systems, particularly when loops are strongly coupled. Consequently, a great deal of insight (or "artistry") is required to apply servo analysis to multi-variable, multi-loop problems.

b. Treatment of pilot workload is cumbersome, especially when the pilot must generate lead in multiple loops.

8

c. Parameters related to the quality of the perceptual environment (e.g., limitations associated with perceptual resolution) are not accounted for. This can be an important consideration, for example, in the landing approach task, where the pilot's ability to perceive and utilize height and sink-rate errors (and, therefore, performance and workload) is a strong function of the physical aspects of the relevant display.

d. At present there is no rule for combining the predicted pilot ratings for individual tasks into a predicted rating for a combined task. For example, the rating associated with a combined longitudinal and lateral-direction flight control task does not bear a consistent relationship to the ratings for the individual subtasks. In order to treat the multi-task (or multi-axis) problem, an analysis scheme is required which accounts for the interaction between multiple tasks performed concurrently.

Building on the ideas of Anderson, staff members of Bolt Beranek and Newman (BBN) Inc., suggested a model-based scheme to overcome some of these limitations.* Attentional workload was defined in terms of a model parameter, and the pilot was assumed to tradeoff between workload and a scalar metric of system performance to minimize the numerical pilot rating. A rating expression, formulated as a function of "workload" and performance, was tested against existing experimental data with encouraging results.

More recently, Hess has described a model-based scheme for predicting pilot ratings that is similar to that proposed by BBN [9]. He suggests an index of performance of the following form:

---

* "A Technique for Predicting Aircraft Handling Qualities as a function of System Performance and Attentional Demand", Technical Memorandum CSD-7, November 1974, Control Systems Department, Bolt Beranek and Newman Inc., Cambridge, Massachusetts.

$$J = E\{\lim_{T \to \infty} \frac{1}{T} \int_0^T \left[\underline{y}'(t) \; \underline{Q}\underline{y}(t) + \underline{u}'(t) \; \underline{R}\underline{u}(t) \right] \, dt$$

$$(2)$$

where $\underline{y}(t)$ is a vector of system variables that the pilot wishes to maintain within acceptable limits, $\underline{u}(t)$ is the set of pilot control inputs, $\underline{Q}$ and $\underline{R}$ are constant weighting matrices, and E represents the operation of statistical expectation. In the case of steady-state tracking tasks, the performance index consists of a weighted sum of mean-squared-error and mean-squared-control terms. The pilot is assumed to adopt control and estimation strategies that minimize this performance index.

Hess proposes a model structure, based on modern (or "optimal") control theory, to allow one to predict the performance index for various flight tasks. Hess' model is a modified implementation of the model originally suggested by Kleinman, Baron, and Levison [10,11]. The latter model forms the basis for the prediction scheme that is the subject of this report.

Hess hypothesizes that pilot opinion rating will be related monotonically to the index of performance if:

1. the index of performance and the pilot-related model parameters required to match the data yield a dynamically representative model of the human controller,

2. the variables included in the performance index are directly observable by the pilot, and

3. weighting coefficients are chosen as the reciprocal of the squares of maximum allowable deviations, where such maxima are consistent with the pilot's perceptions of task requirements.

Hess claims that this prediction scheme accounts for both the physical as well as the mental aspects of pilot workload. Physical workload is defined in terms of the weighted mean-squared

control motions. Mental (i.e., information-processing) workload is related to the rank of the Q weighting matrix appearing in the first term of Eq. (2). That is, mental workload is related to "the number of separate variables whose deviations the pilot considers pertinent to the task and whose deviations contribute to the value of the index of performance" [9].

Hess tested this hypothesis against 19 different configurations covering a range of pilot ratings. Performance "limits" were chosen to match experimental scores, and pilot-related model parameters were chosen partly on the basis of previous results and partly to match observed performance. Pilot ratings could be matched to within $\pm$ 1 rating unit by a linear relationship between pilot rating and the logarithm of the performance index of Equation (2). More recently, Schmidt has used this prediction scheme as the basis for a model-based control design procedure [12].

Although not validated as a reliable predictive tool, Hess' procedure lays the foundation for a scheme that seems to overcome some of the limitations inherent in techniques based on classical servo analysis. The basic form of the performance index is consistent across tasks; the form of the pilot model and nature of loop closures are determined by the optimal pilot model and need not be specified by the user; a scalar metric of workload is provided; factors related to perceptual environment are considered (as shown in the following section of this report); and, in principle, pilot ratings can be predicted for combined longitudinal and lateral axis tasks.

Perhaps the most severe limitation of the optimal-model-based approach, as developed so far, is the requirement to specify numerous task- and pilot-related model parameters. To some extent, the "artistry" in specifying pilot model forms and loop closures for servo-theory models is replaced by the artistry in specifying parameters (especially weighting matrices) of the optimal-control model.

Another limitation, in the opinion of this author, is the lack of a suitable metric for information-processing workload. The metric proposed by Hess (the number of system variables to be regulated) does not appear to add to the rating scheme beyond what is encompassed by the performance index. That is, if workload is to be related to controlled variables that are of concern to the pilot, then only those variables contributing significantly to the performance index will increase pilot workload. Such effects are accounted for by the numeric value of the index itself.

The methodology described in this report builds upon the work of Hess and encompasses a pilot rating prediction scheme based on the optimal-control model for pilot/vehicle performance. Emphasis is placed on the predictive aspects of the procedure, and a rationale is offered for selecting model parameters on the basis of an adequate description of the task and in the absence of experimental data. In addition, a well-defined model parameter is suggested as a potential scalar workload metric for the purposes of predicting pilot opinion ratings.

2. METHODOLOGY

Because pilot opinion is assumed to reflect both pilot workload requirements as well as system performance capabilites, methods for predicting pilot ratings should include consistent and straightforward treatments of workload. Therefore, before proceeding with a description of the proposed rating scheme, let us briefly review the concept of workload as used in this study.

The term "workload" is intended to refer to information-processing -- rather than physical -- activity of the pilot. Specifically, workload is considered synonomous with "attention" in the remainder of this report. Although attention is not defined here in a way that lends itself to direct physical measurement, the pilot model used in the rating prediction scheme does include a parameter that can be related to attention on both theoretical and empirical grounds. Thus, for the purposes of obtaining rating predictions, attention (workload) is an unambiguous and workable concept.

Underlying the use of workload in the prediction scheme is the notion that attention is a voluntary aspect of pilot behavior. That is, we assume that the pilot controls the amount of attention devoted to a particular flight task or subtask, and that, in general, the greater the attention, the better the overall system performance. The workload associated with a task is thus equivalent to the level of attention the pilot decides to devote to the task.

2.1 General Approach

The prediction scheme described in this report is based on the following assumptions:

1. Pilot rating is a function of the flight task.

2. For a given flight task there exist one or more critical subtasks which serve as the primary determinants of pilot rating.

3. Performance requirements are well defined for each critical subtask.

4. Pilot opinion is based partly on the degree to which desired performance is achieved and partly on the information-processing workload associated with the task.

5. A reliable model exists for predicting performance/workload tradeoffs for relevant flight tasks.

These assumptions lead to the procedure diagrammed in Figure 1. The following steps are required for predicting an average pilot rating for a specific situation.

1. Task Definition. Pilot opinion ratings are task dependent. For example, the rating associated with a specific vehicle, relative to other vehicles or other configurations of the same basic airframe, may not be the same in final approach as, say, in high-altitude cruise. Therefore, separate assessments must be made for each flight task of interest.

2. Subtask Definition. Use of the methodology requires a quantitative description of the specific task or subtask for which predictions are to be obtained. For example, if ratings are desired for landing approach, a critical aspect of that task (say, ILS tracking) must be quantified. Task specification requires a linearized description of vehicle dynamics plus a quantitative description of the external environment (e.g., spectral characteristics of the wind gusts if the subtask is path regulation in the presence of zero-mean random turbulence).

3. Define Performance Criteria. Performance criteria must be defined in precise quantitative terms. In order to obtain performance/workload predictions with the pilot/vehicle model used in this procedure, a quadratic performance index of a form similar to that given above in Equation (2) must be specified. The user must specify both the terms to be
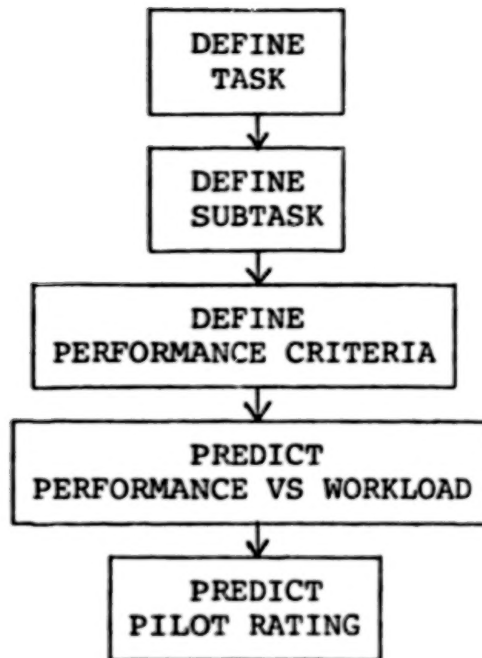
Figure 1. - Procedure for Predicting Pilot Rating.

included in the performance index as well as values for the cost weighting coefficients.

Cost weighting coefficients based on assumed maximum allowable values are suggested. As illustrated in Section 2.3 below, these coefficients are determined partly from the physical constraints of the flight control system, partly from objective performance requirements of the closed-loop system, and partly from pilot preference. The performance criterion used in the rating expression should be a monotonic function of this quadratic performance index.

4. Predict Performance/Workload Tradeoff. The optimal-control pilot/vehicle model described below is used to predict performance as a function of information-processing workload. "Workload" -- considered synonymous with "attention" in the context of the model -- is defined in terms of a model parameter relating to signal/noise characteristics of the human operator.

5. Predict Pilot Rating. The results of the preceding step are used in a rating expression to predict the pilot rating. If experimental data are available for the flight task/subtask of interest, a regression analysis is performed to "calibrate" the independent parameters of the rating expression; in this case, absolute rating predictions are obtained. In the absence of such calibration data, rating parameters are adjusted on the basis of previous results, and rating predictions are interpreted on a relative basis with predictions obtained for other vehicle configurations.

In the remainder of this section we review the pilot/vehicle model and describe alternative expressions for predicting pilot opinion ratings.

2.2 Review of the Pilot/Vehicle Model

The prediction technique described in this report is built around the so-called "optimal-control" model for pilot/vehicle systems. The theoretical foundation for this model has been described in the literature [10, 11], and the model has been validated for both simple laboratory tracking tasks [10,11,13-15] as well as for more complex control situations [16-18]. As discussed above, this model has also been shown to yield good handling qualities predictions [9].

A review of the structure and parameterization of the model is provided in Appendix A; essential features are summarized below.

The model is based on the assumption that the well-motivated, well-trained human operator behaves in a near optimal manner subject to his inherent constraints and limitations. The operator is assumed to adopt strategies of state estimation and control that minimize a "cost function" (or performance index) of the form:

$$J = E\{\lim_{T \to \infty} \frac{1}{T} \int_0^T [\sum_{i=1}^{N_y} q_i y_i^2(t) + \sum_{i=1}^{N_u} (r_i \dot{}^2(t) + g_i \dot{u}^2(t)] \quad dt \quad (3)$$

This expression augments that of Eq. (2) by the addition of cost penalties on the rate of change of control, inclusion of which introduces a first-order lag into the pilot's control strategy. In simple single-variable laboratory tracking tasks, this lag appears to reflect a bandwidth limitation on the part of the human operator; in more realistic control tasks, cost coefficients associated with control rate may be selected to reflect physical limitations on the slew rate of aircraft control surfaces.

Pilot-related limitations reflected in the model include information-processing delay, response bandwidth limitations, response randomness, and limitations related to perception.

Information-processing delay is accounted for by a pure time delay which, for mathematical convenience, is associated with the perceptual process. Generally, a time delay of 0.2 ± 0.05 seconds provides a good match to experimental data. To the extent that control activity is limited by operator response limitations (as opposed to limitations of the physical control system), a good match to experimental data can be obtained by selecting the cost coefficient on control rate to yield a lag time constant of approximately 0.1 seconds. (This lag is not lumped into the pure delay term.)

The "observation noise" and "motor noise" parameters account for response randomness; the former accounts as well for perceptual limitations. The motor noise term is included primarily to reflect limitations in the pilot's knowledge of the response characteristics of his vehicle; a typical value of motor noise is -60 dB, normalized with respect to control-rate variance.[*]

The stochastic portion of the pilot's response ("pilot remnant") is accounted for largely by an observation noise process. Each perceptual variable utilized by the pilot is assumed to be perturbed by a Gaussian white noise process linearly independent of other such noise processes and of external inputs to the system. In the case of a single-variable steady-state tracking task in which perceptual threshold- and saturation-type

---

[*] Various representations of motor noise have been explored during the development of the optimal-control model [19,20]. For the version used in this study, motor noise was represented as a Gaussian white noise process injected in parallel with commanded control rate and normalized with respect to the variance of the commanded control rate.

limitations are negligible, the variance of each observation noise process appears to scale with the variance of the associated perceptual variable. Thus,

$$V_i = \pi P_i \, \sigma^2_{y_i} \qquad (4)$$

where $y_i(t)$ is the ith perceptual variable, $\sigma^2_{y_i}$ is the variance of that variable, P_i is the "observation noise/signal ratio" associated with perception of y_i, and V_i is the autocovariance of the white observation noise processes. This expression can be modified to account for limitations associated with perceptual resolution (see Appendix).

The model is able to reproduce pilot response behavior in a number of simple laboratory tracking tasks with a nearly constant value of noise/signal ratio of about 0.01 (i.e., -20 dB). The consistency of this parameter across tasks and across subject populations suggests that it reflects a basic central-processing (rather than perceptual or motor) limitation, and these results have led to the following model for central attention sharing:

$$P_i = P_o \cdot \frac{1}{f_t} \cdot \frac{1}{f_s} \cdot \frac{1}{f_i} \qquad (5)$$

where f_t is the fraction of attention devoted to the tracking task as a whole, f_s the fraction of attention devoted to subtask "s" (say, longitudinal-axis control), and f_i is the fraction of attention devoted to the ith display in subtask s. P_o is the baseline noise/signal ratio associated with a high-workload single-variable tracking task (typically, -20 dB).

The attention-sharing model of Eq. (5) has a theoretical base [21] and has been validated in a study of multi-axis tracking by Levison, Elkind, and Ward [13], who found that this model yielded accurate predictions of multi-axis system performance. Wewerinke

19

[22] has also obtained generally good agreement between subjective workload assessments and a "workload index" based partly on the model of Eq. (5)*

The following assumptions are usually made with regard to attention-sharing: (1) all attention devoted to the control task is allocated to one or more identifiable subtasks; and (2) all informational quantities within a given task are mutually interfering - that is, information is obtained from one element of the display only at the expense of degrading the information obtained from another display element.** These assumptions lead to the following constraints:

$$\Sigma \ f_i = 1$$

$$\Sigma \ f_s = 1 \tag{6}$$

The remaining attentional variable, f_t, is usually taken as a free parameter of the analysis when predicting performance/workload tradeoffs.

The specific values for f_i and f_s may be chosen (subject to the above constraint) to reflect some hypothesized allocation of attention, or model solutions may be used to find the allocation of attention that yields optimum performance. That is, one may use the model to predict the optimal allocation of attention.

The constraint on f_i given by (6) is made on the assumption that interference is central in origin and not due to overt

---

* Wewerinke's workload index uses both the noise/signal ratio at which the pilot operates, as in the model suggested here, plus the sensitivity of the performance index to fractional changes in this noise ratio.

** No interference is assumed between position and rate information obtained from the same display element.

20

scanning requirements. If large eye movements are necessary, visually obtained information is further degraded because of the apparent loss of perception that occurs immediately before, during, and after each eye movement [23]. This loss is modeled by letting the f_i sum to a value less than unity. Thus, for scanning situations,

$$\Sigma \ f_i = 1 - f_o$$

$$\Sigma \ f_s = 1$$

(7)

where f_o is the fraction of time that perception is "lost", on the average, because of scanning.

The model parameter f_t, representing attention to the control task as a whole, serves as the metric for workload in the proposed handling qualities prediction scheme. Because it is a scalar quantity, it may be used in a straightforward manner to predict handling qualities for multi-variable, multi-axis flight control tasks. Unlike workload metrics used in alternative model-based prediction schemes, the attention parameter defined here has a theoretical as well as empirical basis.

Because the predicted "cost" for a given task increases monotonically with increasing noise/signal ratio, and because noise/signal ratio is related inversely to the attention parameter f_t, cost is a monotonically decreasing function of "workload" as we have defined it here. Thus, if other independent model parameters are kept fixed, tradeoff curves of performance versus workload can be predicted for configurations of interest. As described below, these curves can be further processed to yield predictions of pilot rating.

2.3 Prediction of Pilot Ratings

In keeping with Anderson's philosophy [5], pilot rating is predicted by means of a mathematical expression that includes both

performance and workload effects. In general, "performance" is defined in terms of the performance index of Eq. (3) or some other scalar function of the signal deviations predicted by model analysis. As described above, "workload" is synonymous with the total attention to the task, f_t, which affects performance through the noise/signal ratio.

Best results in this study were obtained through use of a performance metric defined as the joint probability of one or more system variables being outside their respective "limits" (i.e., maximum desirable values).* Although a reasonable match to experimental results could be obtained with the performance index J defined in Eq. (3), the probability metric provided a better match with a simpler rating expression.

The following alternative philosophies have been tested and found to yield good replications of experimentally obtained pilot ratings:

1. Pilot rating is determined by the performance achievable at some particular level of workload;

2. Pilot rating is determined by the workload required to achieve some criterion level of performance;

3. Pilot rating is a continuous function of both performance and workload, and the pilot operates at a workload so as to minimize the numeric value of his rating (i.e., achieve the best rating).

These philosophies were implemented, respectively, by the following rating expressions:

$$R = 1 + 9 \frac{\sigma}{\sigma + \sigma_o} \mid A = A_o \tag{8}$$

---

* Computation of this probability is described in Section 4.3.

$$R = 1 + 9 \frac{A}{A + A_O} \bigg| \sigma = \sigma_O \tag{9}$$

$$R = 10 \left[\frac{\sigma}{\sigma + \sigma_O} + \frac{A}{A + A_O} \right] \tag{10}$$

$$1 \leq R \leq 10$$

where R is the predicted pilot rating on the Cooper-Harper Scale [24]; σ is predicted performance in terms of a probability as defined above, A is the attention model parameter (equivalent to f_t of Eq. (8)), and σ_O and A_O are constants of the rating expressions.* For convenience, we shall refer to these rating expressions as the "performance model", the "attention model", and the "minimum-rating model".

The first two expressions are intended as predictors of rating only, not as predictors of the specific point on the performance-workload tradeoff curve at which the pilot will operate. For example, the performance expression is not intended as a prediction that the pilot will operate at some specific workload level A_O, but rather that the rating will be based on the performance that would be achieved if the pilot were to operate at that level. Similarly, the attention expression -- based on the notion of a "workload index" suggested by Levison, Elkind, and Ward [13] -- is not intended as prediction that the pilot will always work to achieve a fixed criterion level of performance. In other words, the performance and rating expressions imply that rating is based on a "what if" consideration and not on the levels of performance and workload actually achieved.

---

* Numerical values for A_O and σ_O may vary from one expression to the next.

The minimum-rating expression Eq. (10) embodies the notion expressed by Anderson [5] that the pilot trades between performance and workload in such a way as to minimize the rating score. In principle, use of the minimum-rating expression should allow one to predict pilot workload and overall system performance as well as the pilot rating. Nevertheless, even if the models presented herein are highly accurate predictors of performance/workload tradeoffs, there may be practical problems in pinpointing the pilot's "operating point". If there exists a range along the performance-workload curve for which the predicted rating differs negligibly from the minimum, predictions of operating point cannot be more precise than this range of indifference.

3. DATA BASE

The data base used for developing and testing the handling qualities prediction scheme was obtained from two sources: (1) an experimental study performed by Douglas Aircraft Company in 1975, [25], and (2) the results of a questionnaire, submitted during the course of this study, to the test pilots who participated in the Douglas study.

3.1 Description of Experiments

A manned simulation study was conducted by Douglas Aircraft Company to explore the applicability of various handling qualities criteria to longitudinal flying qualities of large transport aircraft in the landing approach. Criteria that were evaluated included several vehicle-centered criteria from MIL-F-8785B [1], vehicle-centered criteria from other sources [25], and a pitch tracking criterion involving a closed-loop pilot model [2]. This study is described in detail by Rickard [25]; a summary of the experiments is given below.

The Douglas study explored a total of 42 vehicle configurations. The first group of 26 configurations were obtained by selecting stability derivatives typical of wide-body aircraft and either varying the simulated cg location from far forward to far aft of the neutral point, or by varying a single stability derivative. Configurations of the second group were obtained by specifying vehicle frequency-response characteristics and then solving for the stability derivatives. All handling-qualities variations were confined to the longitudinal control axis; lateral-directional aircraft parameters were kept fixed throughout the experiment to provide response characteristics typical of a wide body transport.

Five Douglas test pilots performed evaluations of these configurations on a six-degree-of-freedom moving-base simulator. Each evaluation typically consisted of two ILS approaches: the first performed in the absence of simulated atmospheric

25

disturbances, the second in the presence of simulated zero-mean turbulence. Approach was initiated at a range of 13.7 kilometers (7.4 n. mi.) from runway threshold at an altitude of 457 meters (1500 feet) on the extended runway centerline. The 3-degree glide slope was intercepted at a range of about 8.7 kilometers (4.7 n. mi.); the pilot flew down the glide slope relying on ILS instrumentation for path information to an altitude of about 61 meters (700 feet), at which point the pilot transitioned to a visual display for flare and touchdown.

The test pilots were encouraged, in general, to perform maneuvers that would aid in their evaluations (e.g., intentionally impose and then eliminate a path or attitude error), but no specific set of maneuvers was required. A single Cooper-Harper rating was given by each pilot for the pair of still-air and turbulent-air approaches for each configuration. Some configurations were evaluated more than once by some of the test pilots. Evaluations were performed on the basis of approach performance only; flare and touchdown characteristics were not considered.

3.2 Configurations Explored in the BBN Study

3.2.1 Airframe Response Characteristics

The rating expressions described in Eq. (8) - (10) were tested against eight configurations selected from the first group used in the Douglas study. These configurations were chosen to span a range of pilot ratings as well as a range of handling qualities problems. Modal characteristics for the configurations explored in the BBN study are given in Table 1.

3.2.2 Control and Display Subsystems

The Douglas manned simulation study was performed with the following actuator and engine-response characteristics:*

---

* W. W. Rickard, personal communication.

TABLE 1

Configuration Characteristics

v = 140 kts γ = -3° wt = 1,560,000 N

Config.* Number	ω_{sp}	ξ_{sp}	ω_{ph}	ξ_{ph}	n/α	$d\gamma/dV$	$1/T_{\theta 1}$ or $[\omega]$	$1/T_{\theta 2}$ or $[\xi]$
1	0.846	0.628	0.186	0.072	3.80	-0.040	-0.084	-0.506
3	(-0.633)	(-0.307)	0.086	0.318	4.14	-0.049	-0.082	-0.556
4	(-0.811)	(+0.090)	0.200	0.636	4.20	-0.051	-0.082	-0.564
5	(-0.909)	(+0.158)	0.210	0.479	4.24	-0.053	-0.082	-0.568
8	0.811	0.662	0.194	0.041	4.04	+0.339	+0.041	-0.631
15	(-0.991)	(+0.225)	0.211	0.388	4.29	-0.055	-0.082	-0.575
16	(-1.061)	(+0.291)	0.210	0.331	4.35	-0.057	-0.082	-0.583
21	0.441	0.665	0.170	0.043	1.05	+0.285	[0.149]	[0.676]

ω_{sp} = short-period natural frequency, rad/sec

ζ_{sp} = short-period damping ratio

ω_{ph} = phugoid natural frequency, rad/sec

ζ_{ph} = phugoid damping ratio

n/α = normal acceleration per unit angle of attack, g/rad

$d\gamma/dV$ = path angle change per speed change, deg/kt

T_{θ} = numerator time constant, sec

(   ) signifies first-order factor

*Configuration number of the Douglas Study [25].

$$\frac{\delta_e}{F_{cc}} = -0.056 \cdot \frac{850}{s^2 + 45s + 850} \cdot \frac{2500}{(s + 25)^2} \qquad (11)$$

$$\frac{\delta_T}{\delta_{T_c}} = \frac{2.5}{s + 2.5} \qquad (12)$$

where δ_e/F_{cc} represents degrees of elevator deflection per Newton of control force. As discussed in Section 4.2, these dynamics were simplified for model analysis.

The test pilots were assumed to utilize the ILS instrument, attitude indicator, and airspeed indicator as their primary displays during the instrument-flight portion of the simulated approach.

Glide slope error was indicated by displacement of a vertically-moving indicator with respect to a marker. Calibration markings ("dots") were located above and below the zero reference at separations corresponding to 0.35 degrees of glideslope error. Separation between dots was .874 cm, leading to an effective display gain of 2.5 cm of display deflection per degree glideslope error.

Pitch attitude changes were indicated by motion of the attitude sphere relative to a stationary reference bar; calibration markings were located at intervals corresponding to increments of 2.5 degree pitch attitude. For small deviations from trim, display movement was considered to be linearly related to simulated pitch attitude with a display gain of .163 cm indicator displacement per degree of pitch.

Airspeed was indicated by a rotary pointer that moved relative to a stationary background. Graduations were linearly spaced over the range of 100 to 250 kts; in this range, a 90-degree displacement of the airspeed indicator corresponded to a

change of 70 kts. The pointer was about 3.3 cm in length, providing a display gain of about .074 cm display deflection per knot, or .072 cm/knot.

3.2.3 Random Turbulence

Zero-mean turbulence was simulated in the three linear and three rotational degrees of freedom in the Douglas study. The following turbulence models, based on models suggested in the flying qualities specifications [1], were used to provide disturbances to longitudinal-axis variables:

$$u_g = \sigma_u \sqrt{\frac{2 L_u}{\pi U_o}} \cdot \frac{1}{1 + \frac{L_u}{U_o} s} \cdot N_u \tag{13}$$

$$= \frac{K_u}{s + \alpha_u} \quad W_u$$

$$\tag{14}$$

$$w_g = \sigma_w \sqrt{\frac{L_w}{\pi U_o}} \cdot \frac{1 + \frac{\sqrt{3} L_w}{U_o} s}{(1 + \frac{L_w}{U_o} s)^2} \quad N_w$$

$$= K_w \quad \frac{s + \alpha_{w1}}{(s + \alpha_{w2})^2} \quad W_w$$

$$q_g = 57.3 \cdot \frac{\frac{1}{U_o} s}{1 + \frac{4b}{U_o} s} \quad w_g \tag{15}$$

$$= \frac{K_q}{s + \alpha_q} \quad \dot{w}_q$$

29

where u_g, w_g, and q_g are gusts acting along the forward, vertical, and pitch body axes; N_u and N_w and unity white noise processes; W_u and W_w are white noise processes having variances equal to the u- and w-gust variances, and

$$L_u = \begin{cases} 44.2 + 1.305 \; h, \; m & 0 \leq h \leq 305 \; m \\ 442 \; m & h \geq 305 \; m \end{cases}$$

$$L_w = \begin{cases} h \; m & 0 \leq h \leq 305 \; m \\ 305 \; m & h \geq 305 \; m \end{cases}$$

$$b = 50.4 \; m$$

$$U_o = 72.1 \; m/sec$$

Rms gust level σ_w and σ_w were fixed at 2.4 meters/sec (7.8 ft/sec and 2.0 meters respectively, for the duration of the approach. Other parameters of the turbulence model are quantified in Table 2. These values correspond to an altitude of 305 meters (1000 ft), the altitude for which steady-state model analysis was

Table 2

Parameter Values for the Turbulence Model

Height = 305 m

Parameter	Value
a_u	0.163
K_u	0.571
a_{w1}	0.136
a_{w2}	0.236
K_w	0.842
a_q	1.12
K_q	0.272

performed in the BBN Study.

These models were simplified somewhat for steady-state model analysis as described in Section 4.2.

3.3 Performance Requirements

Application of the prediction scheme in Section 2 requires, first, that one or more specific subtasks be selected for analysis and, second, that performance requirements be specified for each subtask. To obtain this information, a questionnaire was prepared by BBN and administered by Douglas personnel to 4 of the 5 test pilots that had participated in the 1975 manned simulation study. Through this questionnaire the pilots were requested to (1) state whether or not pilot ratings were determined primarily by longitudinal handling characteristics; (2) specify whether ratings were based mainly on the instrument-flight or visual-flight portions of the approach; (3) specify, in order of priority, the subtasks that were important determinants of pilot rating; and (4) specify in as quantitative manner as possible the "desired" and "acceptable" levels of performance for each subtask. A sample of the questionnaire is provided as an Appendix B to this report.

All four pilots agreed that lateral-directional handling qualities were quite satisfactory and that pilot ratings were influenced primarily by longitudinal handling characteristics. They all stated that the instrument-flight phase was more important in determining ratings.

All subjects indicated at least three subtasks as important determinants of pilot rating. Relative importance of these subtasks for the subject population as a whole was determined from "priority scores", computed by assigning 5 "points" to an item receiving highest priority, 4 points to the next priority, and so on to 1 point for tasks ranked fifth or more in the list. Priority scores for each task are shown in Table 3, along with the total score obtained by summing across pilots.

Table 3 shows that ratings were largely determined by the ability of the pilot to regulate path error. Highest priority was given to tasks involving transient maneuvering (glide-slope

Table 3

Priority Scores for Important Subtasks

| Subtask | LB | Priority Score | | | Total |
		BM	JM	AT	
Glide-Slope Capture	5	5	4	4	18
Glide-Slope Tracking	4	-	2	5	11
Recover from Glide Slope Mistrim	3	3	1	4	11
Altitude Station-Keeping	-	4	5	-	9
Open-loop Response	-	2	3	3	8
Recover from Airspeed Mistrim	2	-	1	-	3
Recover from Pitch Mistrim	-	-	1	-	1

capture, correcting self-induced height error). Next in importance were tasks requiring continuous regulation of height error (altitude station-keeping prior to glide slope acquisition, post-acquisition glide-slope tracking). Open-loop response and correction of pitch and airspeed mistrim were of substantially less importance overall in terms of influencing pilot opinion.

Obtaining quantitative comments related to performance requirements was considerably more difficult than anticipated. Only two of the four pilots provided quantitative responses, and only one of these (Subject JM) differentiated between "desired" and "acceptable" performance.* Performance requirements indicated by Subject JM for tasks requiring continuous regulation are given in Table 4.

3.4 Pilot Ratings

Mean and standard deviations of the pilot ratings obtained in the Douglas study are given in Table 5, along with handling qualities levels as determined from two of the vehicle-centered criteria considered by Rickard [25]. Rating statistics were

---

* To aid the pilot in making this distinction, "adequate" performance was defined in the questionnaire as corresponding to the boundary between a rating of 6 and 7, whereas "desired" performance was to be associated with a rating of 1. See Appendix B.

Table 4

Subjective Performance Requirements

	JM				BM
	Alt. Regulation		G-S Tracking		G-S Tracking
	Desired	Acceptable	Desired	Acceptable	
Height Error	± 30 ft.	± 100 ft.	± 1/4 dot	± 1/4 dot at 200' ± 1 dot at 1100'	± 1/4 dot
Sinkrate Error (ft/sec)	± 200	± 500	± 200	± 530	-
Airspeed Error (knots)	± 5	± 10	± 2	± 5	- 5, + 10

derived by first averaging multiple ratings (where such existed) for each pilot for each configuration, and then using these averages to compute a mean and standard deviation across subjects

Table 5

Pilot Opinion Ratings

Config.	Pilot Rating Mean	SD	$d\gamma/dV$ Level	ω_{sp} vs. n/α Level	Static Stability
1	2.5	1.5	1	1	Yes
3	4.3	2.3	1	4	Yes
4	4.2	2.1	1	4	No
5	5.3	1.6	1	4	No
8	8.3	2.1	4	1	Yes
15	6.7	1.5	1	4	No
16	7.7	2.5	1	4	No
21	6.2	3.5	4	2	Yes

Mean rating for 5 pilots, configurations 1,2,4,5
Mean rating for 3 pilots, configurations 8,15,16,21

for each configuration.[*]

Table 5 shows both a wide spread of average pilot ratings across the configurations explored in the BBN study as well as a variety of handling qualities problems. The short-period response criterion predicts adverse handling qualities for five of the configurations -- four of which exhibit static instability. Two of the remaining configurations, on the other hand, exhibit adverse flight path stability ($d\gamma/dV$).

---

[*] Ratings shown for configurations 4 and 5 differ slightly from those presented by Rickard, who apparently computed the mean of all ratings pertaining to a given configuration regardless of the number of evaluations per pilot.

4. TEST OF METHODOLOGY

The steps to be followed when obtaining a prediction of handling qualities -- diagrammed previously in Figure 1 -- are repeated in Figure 2 for convenience. In addition to the five steps designated here, there is usually the additional process of problem simplification: computational requirements are reduced through the use of reduced-order models for various system components (airframe dynamics, turbulence model, control system dynamics, etc.) wherever possible without materially compromising predictive accuracy.

The results of applying the methodology proposed in Section 2 to the data base described in Section 3 are described below. Discussion is organized as follows: (1) parameters of the prediction scheme,(2) problem definition, including specification of independent model parameters; (3) problem simplification; (4) prediction of performance as a function of attentional workload; and (5) prediction of pilot ratings, and (6) discussion of results.

4.1 Parameters of the Prediction Scheme

A number of model parameters must be quantified in order to predict pilot ratings. These parameters fall into four basic categories: (a) cost coefficients, (b) display-related limitations, (c) attentional allocation, and (d) parameters of the rating expression(s).   The number of parameters to be quantified depends on the size of the problem. A cost coefficient will be associated with each system variable of particular concern to the pilot plus two for each control variable (one for control force or displacement, one for control rate). An effective perceptual threshold is associated with each perceptual variable assumed to be utilized by the pilot; in addition, a "residual noise" is associated with each displacement variable for which an explicit null reference is absent (usually pitch). Also, allocation of attention among these displays must be assumed or calculated.

34

Finally, the two "free" parameters of the rating expression must be quantified.

Seventeen independent model parameters were quantified for the problem selected to test the proposed methodology. As the following discussion demonstrates, parameters falling into the first three groups listed above (fifteen in all) were defined largely on the basis of task analysis, tempered by some engineering judgement. Once selected, these parameter values were held fixed throughout the analysis; only the two parameters of the rating expression were adjusted to match experimental data.

4.2 Problem Definition

As already indicated, the methodology described in this report was applied to the general flight task of final approach (exclusive of landing). On the basis of the results of the questionnaire administered to the Douglas test pilots, continuous glide-slope tracking in turbulence and recovery from glide-slope offset were initially selected as the specific subtasks to be explored. Although preliminary exploration of the latter (transient) task was performed, resources permitted a complete analysis of only the continuous tracking task. Therefore, presentation of results is confined to tests based on the continuous tracking task.

Although continuous in nature, glide-slope tracking following capture is not, strictly speaking, a steady-state task because of time variations in various task parameters. As indicated in Equations (13) to (15), turbulence bandwidth changes with altitude. Because path control (in terms of linear, rather than angular, measure) becomes more important as the touchdown point is approached, the pilot's control strategy will change as a function of distance-to-go [18,26]. In addition, since the ILS instrument displays path error in terms of angular deviation, the effective display gain (cm of indicator deflection per meter of height error) also varies with range. Nevertheless, because these time
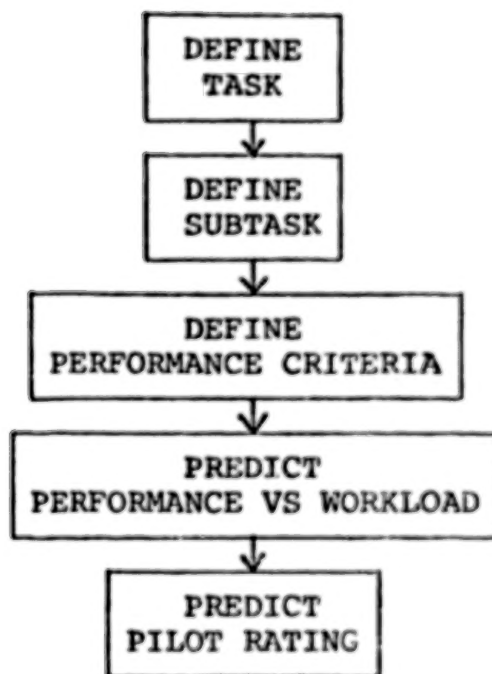
Figure 2. - Procedure for Predicting Pilot Rating.

variations are slow compared to the time constants of important system variables, piecewise-steady-state analysis can yield meaningful predictions of pilot/vehicle performance at various points along the glide path.

A "frozen-point" analysis was performed at a simulated altitude of 305 meters (1000 feet). Parameters of the turbulence model appropriate to this altitude -- indicated in Table 2 -- were chosen for this analysis. Remaining system parameters, which were not dependent on range or altitude, were initially selected to mimic the experimental math models presented in Section 3.

"Maximum allowable" values, or "limits", were determined as follows. A height limit of 35.7 meters (117 feet), corresponding to a glide-slope error of 1 dot, was chosen on the basis of the "acceptable" height error indicated by one of the test pilots (see Table 4). Also based on the results presented in Table 4, a limit of 10 kts, or 5.14 meters/sec (16.9 ft/sec), was associated with airspeed error.

No limits (i.e., no terms in the quadratic "performance index") were associated with either sinkrate error or attitude variables. Although one subject specified limitations on sinkrate error, this variable was omitted from the performance index to prevent overemphasis on height-related variables. Penalties on attitude variables were omitted because no limitations on such variables were specified by the test pilots.

Despite the lack of explicit performance penalities on attitude variables, reasonable model predictions were obtainable through appropriate limits associated with control-related variables. Limits on elevator deflection and rate were based on physical limitations of the control actuator systems. The actual limit on elevator deflection was $\pm$ 20 degrees, or 356 newtons (80 pounds) stick force as determined from the actuator model of Eq. (11). The "limit" used for determining a performance penalty was half this value, or 178 newtons (40 pounds). Similarly, the

physical limitation on elevator slew rate -- 30 deg/sec -- yielded a stick-rate limit of 267 newtons/sec (60 pounds/sec).

The limit associated with thrust was set equal to the available thrust deflection from trim: 95,700 newtons (21,500 pounds). The limit associated with rate of change of thrust was based on the assumption that the pilot would not make continuous wide-band throttle inputs. Therefore, a limit of 47,900 newtons/sec (10,750 pounds/sec) was selected to induce a control-related lag time constant of about 2 seconds.

The pilots were assumed to make longitudinal-axis flight-control inputs primarily on the basis of perceptual information obtained from the ILS, attitude, and airspeed instruments. Rate information was also assumed to be obtained from the ILS and attitude indicators. Thus, the "display vector" assumed for model analysis consisted of height, sinkrate, pitch, pitch rate, and airspeed errors.

Attention was assumed to be divided equally between the ILS, attitude, and airspeed instruments;* no attention-sharing penalties were considered between displacement and rate information from the same physical display. On the basis of analysis performed in a previous analytic study of landing approach [18], 34% of the pilot's attention was assumed to be "lost" because of large eye movements required to scan the flight-control instruments. Thus, fractional attentions of 0.22 were associated with the ILS, attitude, and airspeed displays.

---

* To be entirely consistent with the notion of optimal pilot response behavior, an allocation of attention should be determined that minimizes the performance index. Previous studies have shown, however, that an equal allocation of attention among essential display variables yields model predictions very close to those obtained with optimal attention sharing [28]. Therefore, to simplify the analysis, uniform attention-sharing was assumed.

Effective perceptual "thresholds" were computed as indicated below (see Appendix A):

$$a = \frac{D}{G} \cdot \frac{a_o}{57.3} \tag{16}$$

where "a" is the threshold in problem units, a_o is the perceptual threshold in visual units, D is the viewing distance, and G is the display gain in inches of display deflection per unit of "error" (problem units).

The viewing distance for the Douglas simulation experiments was 0.80 meters. Perceptual thresholds of 0.05 degree visual arc and 0.2 arc degree/second were assumed on the basis of a previous laboratory study [13]. These parameter values, along with the effective display gains given in Section 3.2, yield the effective thresholds shown in Table 6. Also shown in Table 6 are the cost weighting coefficients and a "residual noise" associated with perception of pitch error. Cost coefficients were computed as the reciprocals of the squares of associated "limits" (also shown in the table); the rms residual noise was taken as distance of the trim pitch angle (3°) from the nearest calibration marking (2.5^o).

4.3 Problem Simplification

The following simplifications were made to the mathematical model of the flight-control task to improve computational efficiency:

1. Control actuator dynamics of Eq. (11) were represented by a pure delay of 0.09 sec, which was added to the pilot's time delay when predicting closed-loop response.

2. Engine response dynamics of Eq. (2) were ignored. (The large response lag imposed by the selection of a cost coefficient on rate of change of throttle more than accounted for engine response lag.)

Table 6

Display- and Performance-Related Model Parameters

Variable	Limit	Cost Coefficient	Threshold	Residual Noise	Relative Attention
h	35.7	7.85 E-04	2.8	0	.22
$\dot{h}$	-	-	11.	0	.22
θ	-	-	.43	0.5	.22
q	-	-	1.72	0	.22
u_i	5.15	3.77 E-02	0.58	0	.22
δ_e	40	6.25 E-04	-	-	-
$\dot{\delta}_e$	60	2.78 E-04	-	-	-
δ_t	95,700	1.09 E-10	-	-	-
$\dot{\delta}_t$	47,800	4.38 E-10	-	-	-

h = altitude error, meter

θ = pitch change, degrees

q = pitch rate, degrees/second

u_i = airspeed relative to moving air mass, meter/second

δ_e = force on the control column, Newton

δ_t = thrust deviation from trim, Newton

3. The pilot's use of motion cues was ignored, thereby saving two state variables required for a second-order model of simulator platform response dynamics.

4. The following second-order approximation to the turbulence model of Eqs. (13) to (15) was used:

$$u_g = \frac{.571}{s + .163} \ W_u$$

$$w_g = \frac{.824}{s + .340} \ W_w \quad\quad (17)$$

$$\dot{w}_g = \frac{.824s}{s + .340} \ W_w$$

These simplifications resulted in a saving of 9 state variables required to model the task (18 state variables for full-state modelling versus 9 state variables for the reduced-order representation). Because computational time is approximately proportional to the cube of the number of state variables, computational time was thereby reduced by 7/8!

The effects of model simplification on predicted performance were explored for the baseline configuration (Configuration 1 of the Douglas study). Cost and display-related model parameters were selected as shown in Table 6. In addition, the pilot's time delay was set at 0.29 seconds (0.2 sec for nominal pilot delay plus 0.09 sec for control-actuator response), "pseudo" motor noise/signal ratio was fixed at -60 dB,* and the attentional parameters f_s and f_t of Eq. (5) were set to unity.

Table 7 provides a comparison of model predictions made with and without the simplifications. Predictions of rms performance scores are given for the various display quantities assumed available to the pilot and for control-related variables. Model configuration "A" corresponds to the reduced-order 9-state representation, configuration B includes the full 4-state

* See Levison, Baron, and Junker [19] for a description of the treatment of motor noise.

Table 7

Effects of Model Simplification on Predicted rms Response

Variable	Model Configuration A	B	C
h	8.51	7.15	7.56
$\bar{h}$	1.60	1.53	1.49
θ	1.62	1.75	1.63
q	.625	.850	.610
u_i	2.08	2.11	2.07
δ_e	4.54	4.70	4.30
$\dot{\delta}_e$	3.23	3.22	2.81
δ_t	1220	1040	1220
$\dot{\delta}_t$	271	241	2720

A: Simplified Model
B: Full Turbulence Model, Engine Response
C: Simplified Model with Motion Cues

turbulence model plus the model for engine response, and configuration C corresponds to the simplified model with the addition of motion cues.

To include the effects of motion cues in the analysis, the pilot was assumed to obtain n_z information from the moving cab. Cab dynamics were represented by a second-order Butterworth filter having a break frequency at 7.5 rad/sec; the output of this filter was added to the pilot's "display vector". An attention f_i of 0.66 was associated with perception of this quantity to reflect the assumption of no attention-sharing requirements between visual and motion cues [19].

Table 7 shows that most of the predicted scores for model configurations B and C were within 15% of the scores predicted for the reduced-order configuration A. The greatest exception to this rule was the predicted pitch-rate score, which was 35% greater for configuration B than for A. However, predicted height and airspeed errors -- the scores that contributed most to the performance index -- varied by less than 15%; therefore, the simplified model structure was considered to yield representative predictions of overall system performance as defined below.

4.4 Prediction of Performance/Workload Tradeoffs

Performance/workload tradeoffs were predicted for each of the eight configurations defined in Section 3. For purposes of predicting handling qualities, "performance" was defined as the probability that one or more system variables exceeds maximum allowable values. To obtain an approximation to this joint probability, system variables were treated as independent Gaussian variables, and the probability was computed as

$$Pr = 1 - \prod_i (1 - Pr_i) \qquad (18)$$

where Pr_i is the probability that the ith variable of interest will lie outside its prescribed boundary, and Pr is the probability that at least one such variable is out of bounds. The probability Pr_i was readily computed from the predicted variance of the ith system variable. (Since we considered steady-state conditions, all variables were assumed to be zero-mean Gaussian processes.)

"Workload" was represented in the analysis by the attentional variable f_t of the submodel shown in Eq. (5). The variable f_s was set to unity (i.e., we ignored the attentional requirement of the lateral-axis task), and the f_i were adjusted to reflect attention-sharing as shown in Table 6. A noise/signal ratio $P_o = 0.01$ was associated with a relative attention of unity. Thus, variations in attentional workload were reflected by changes in the noise/signal ratios according to Eq. (5).

Analysis was initially performed under the assumption that the throttle was not used in a continuous manner for flight-path control.[*] For this phase of the analysis, no limit (and, hence, no cost penalty) was assigned to airspeed, and the (mathematical) "pilot" was assumed to use only the elevator control. Performance was computed with respect to limits on height error, stick force, and rate of change of stick force as given in Table 6.

Figure 3 shows that the performance/workload curves fall into two general categories: the curves for configurations 8 and 21 for which performance is relatively insensitive to attention, and the remaining six conditions that show a strong dependency. When plotted on a logarithmic basis, as is done in Figure 3, these curves show a nearly linear relationship between performance and

---

[*] This assumption reflected the judgement of this author and of professional colleagues involved in this study. Questions regarding use of the throttle in the Douglas simulation study were not part of the questionnaire given to the Douglas test pilots.
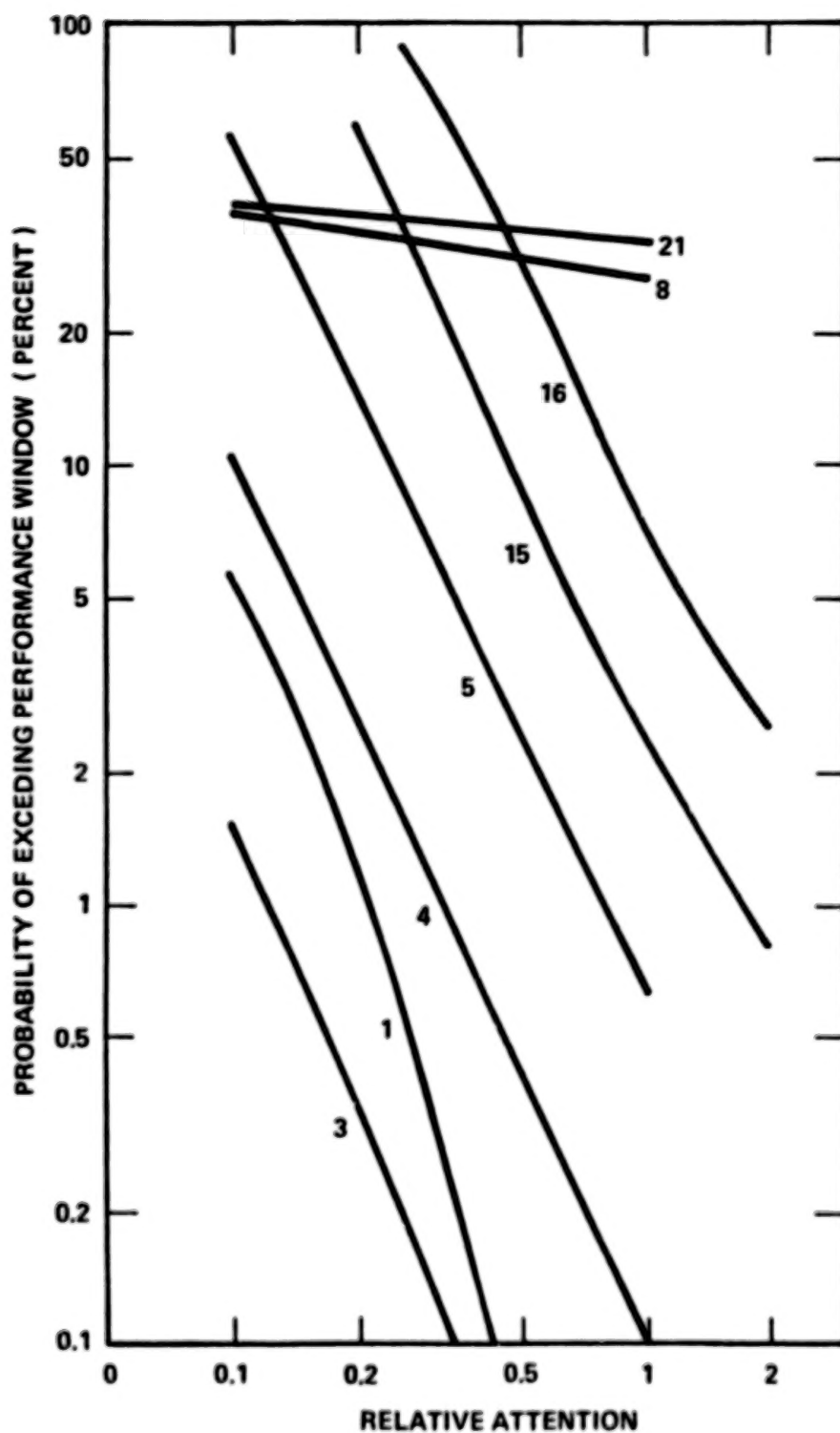
44

Figure 3. - Predictions of Performance versus Workload: No Throttle Control.

workload.[*]

The group of six curves shows an ordering largely consistent with the pilot ratings given in Table 5. The more favorably rated configurations (1,3,4) show the lowest predicted probabilities of exceeding their limits; those corresponding to the least favorably rated conditions (15,16) show greater predicted errors; and the curve for condition 5, which received an intermediate rating, falls between the above two groupings.

Performance scores for conditions 8 and 21 are not consistent with the results of the manned simulation. Predicted probabilities of exceeding performance limits are greater than 25% even at relatively high workload levels; yet, the test pilots were seldom unable to complete a successful simulated landing.[**]

Because of the anomalous predictions obtained for configurations 8 and 21 with throttle control omitted, the above analysis was repeated with inclusion of throttle control and a requirement to maintain airspeed. Performance requirements were as shown in Table 6; other model parameters remained constant.

Figure 4 shows considerable improvement in predicted performance for configurations 8 and 21, especially for high workload. Though not large in absolute terms, predicted probabilities for most of the remaining conditions at high workload are substantially greater than predicted in the previous analysis. Apparently, the requirement to maintain airspeed places a "floor" on performance; factors other than attention (such as

---

[*] Some of the curves of Figure 3-5 are shown for attention levels greater than unity. Values of "attention" are relative to that inferred for data obtained in a standardized laboratory tracking task. Thus, unity attention is intended as a benchmark level of workload and does not necessarily relate to maximum effort or capability.
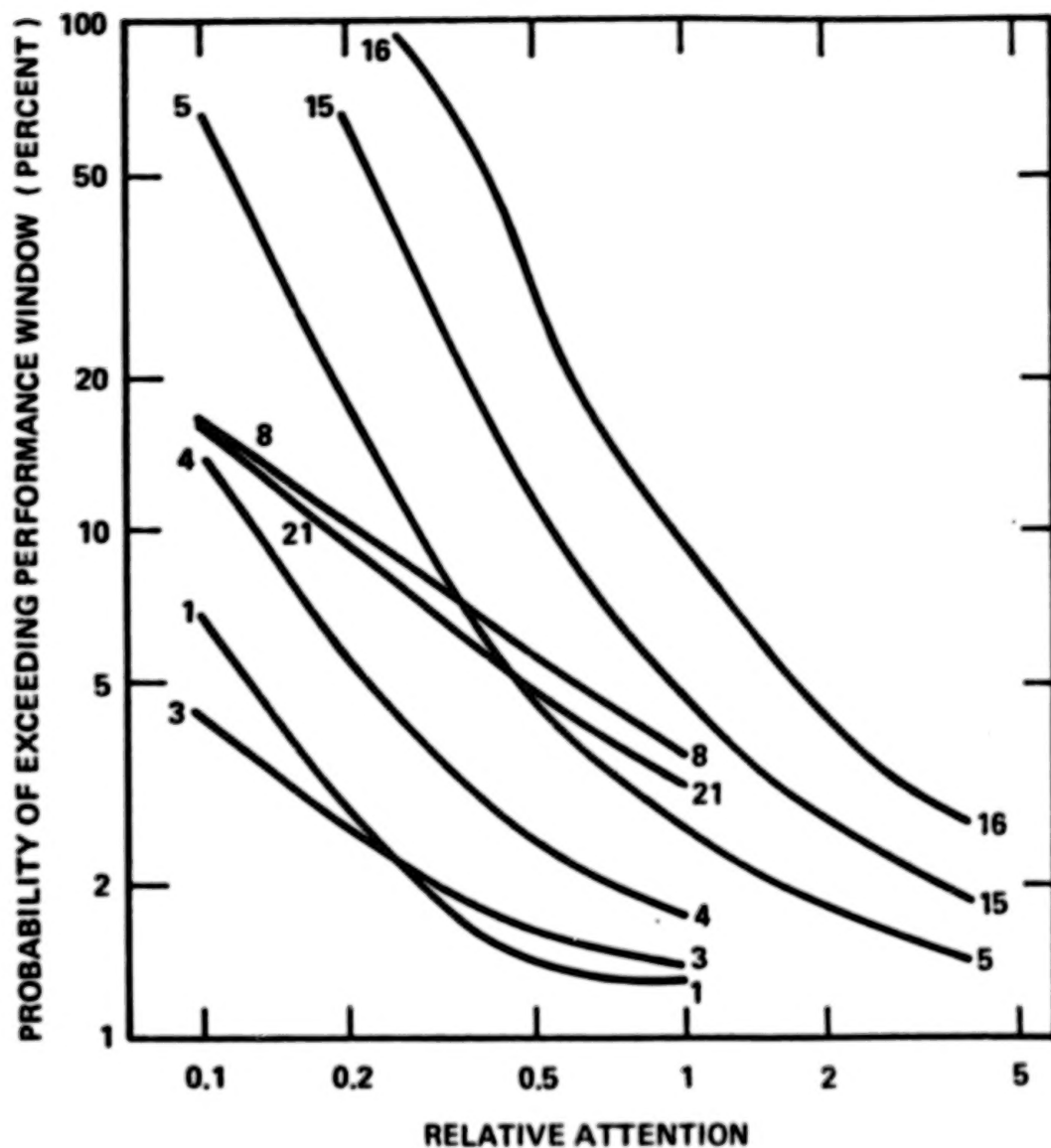
[**] W. W. Rickard, personal communication.

Figure 4. - Prediction of Performance versus
Workload: Throttle Control.

gust and vehicle response bandwidths) limit the capability for reducing airspeed error below a certain amount.

Except for configuration 8, the ordering of the performance/workload curves is consistent with the ordering of the pilot ratings. For attentions of 0.5 and greater, predicted performance for the remaining seven configurations follows the trend of the ratings. Operation on these results to yield predicted pilot ratings is discussed below.

4.5 Predicted Ratings

The three rating expressions presented in Section 2 were applied to the performance/workload tradeoff curves to provide a test of the proposed methodology. These rating expressions are repeated below, along with values assigned to the independent parameters of each expression:

1. The "performance model":

$$R = 1 + 9 \frac{\sigma}{\sigma + \sigma_o} \Big| A = A_o$$

$$\sigma_o = 0.053$$
$$A_o = 0.5$$

2. The "attention model":

$$R = 1 + 9 \frac{A}{A + A_o} \Big| \sigma = \sigma_o$$

$$A_o = 0.47$$
$$\sigma_o = 0.05$$

3. The "minimum-rating "model":

$$R = 10 \left(\frac{\sigma}{\sigma + \sigma_o} + \frac{A}{A + A_o} \right)$$

$$1 \le R \le 10$$
$$\sigma_o = 0.10$$
$$A_o = 2.2$$

The value for A_o of the performance model was chosen to represent a moderate-to-high workload level, and the corresponding value for σ_o was found through a regression procedure that minimized the following "loss function":

$$E = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\hat{R}_i - \bar{R}_i}{\sigma_i} \right)^2 \right\}^{\frac{1}{2}}$$

<div align="right">(19)</div>

where $\hat{R}_i$ = predicted rating, $\bar{R}_i$ = mean experimental rating, σ_o = standard deviation of experimental rating, and E = modeling error.

The value for σ_0 of the attention model was selected to represent a moderate-to-stringent performance requirement, and the value for A_0 was found through the regression procedure described above.

Because of the lack of a tractable analytic expression relating performance to workload, the parameters σ_0 and A_0 of the minimum-rating model were not found through a computerized regression analysis. Rather, pairs of integers were explored on a trial-and-error basis to provide a good match to experimental pilot ratings. The predicted (minimum) rating for a given configuration was obtained by superimposing the predicted performance/workload tradeoff curve (Figure 4) on the curve of constant rating, shown in Figure 5.

Because of the difficulty in matching the predicted pilot rating of Configuration 8, ratings for this configuration were omitted from all three regression analyses. Predicted ratings for the three schemes are compared with experimental ratings for all eight configurations in Table 8.

Figure 6 provides a graphical comparison of predicted versus experimental pilot ratings for the three rating expressions. Dashed lines indicate boundaries of $\pm$ 1 rating unit. The three rating schemes performed about equally well on the average and were able to match 6 of the 8 experimental ratings to within a rating unit. The configuration matched least well was Configuration 8, which was omitted from the regression analyses.
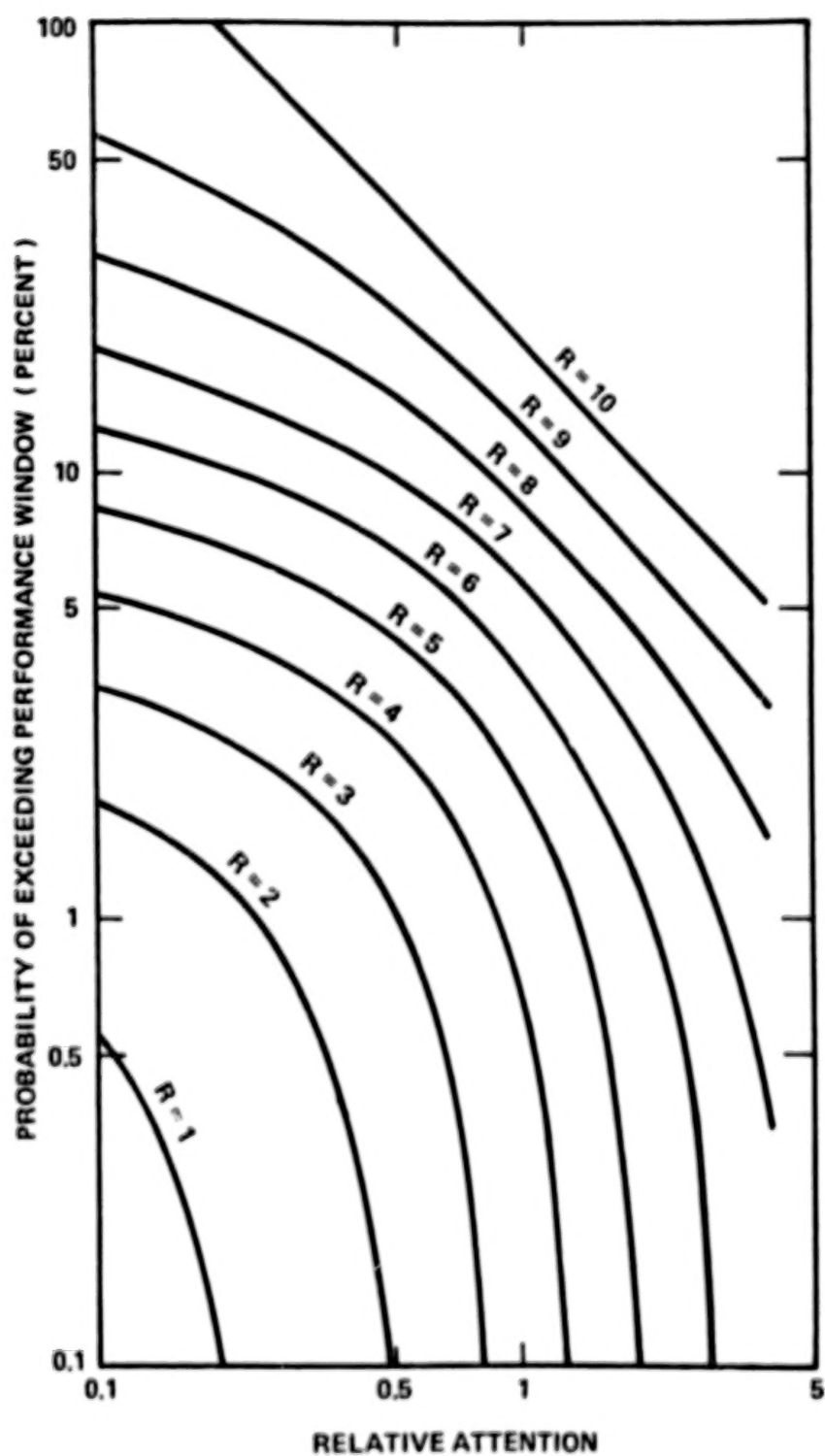
Figure 5.   Curves of Constant Rating for the
"Minimum Rating" Model

$\sigma_o = 10\%$ $A_o = 2$

Table 8

Predicted and Experimental Pilot Ratings

Pilot Ratings

Config.	Exptl.	A	B	C
1	2.5	2.9	2.8	2.9
3	4.3	3.1	2.6	3.0
4	4.2	3.8	3.9	3.9
5	5.3	5.2	5.5	5.1
8	8.3	5.5	6.0	5.7
15	6.7	7.2	7.0	6.6
16	7.7	8.7	8.0	8.0
21	6.2	5.3	5.6	5.4

A: Performance Model
B: Attention Model
C: Minimum Rating Model

Prediction errors may be compared against the variability of the experimental data in Figure 7. Experimental ratings are indicated by filled circles, with brackets to indicate $\pm$ standard deviation; open symbols indicate predictions obtained with the three rating expressions.

Except for Configuration 8, predicted ratings are within one standard deviation of the experimental mean. Even for the worst case, the prediction error is well within two standard deviations of the mean. Thus, the reliability of the predicted ratings is commensurate with the reliability of the experimental data.

To compare the results of these model-based schemes with other handling qualities critera, we adopt the definitions used by Rickard to relate pilot ratings to handling qualities levels. These definitions are given in Table 9. The use of half-levels is adopted so that a predicted rating that is just across the level boundary from a corresponding experimental rating is not considered to be in error by a full unit.

Table 10 compares handling qualities levels obtained experimentally with levels obtained from various prediction schemes. (Predicted levels for other than the three schemes described in this report were obtained from Rickard [25]). The "error" shown in the last column is the accumulated prediction error across the eight configurations in terms of half rating units.
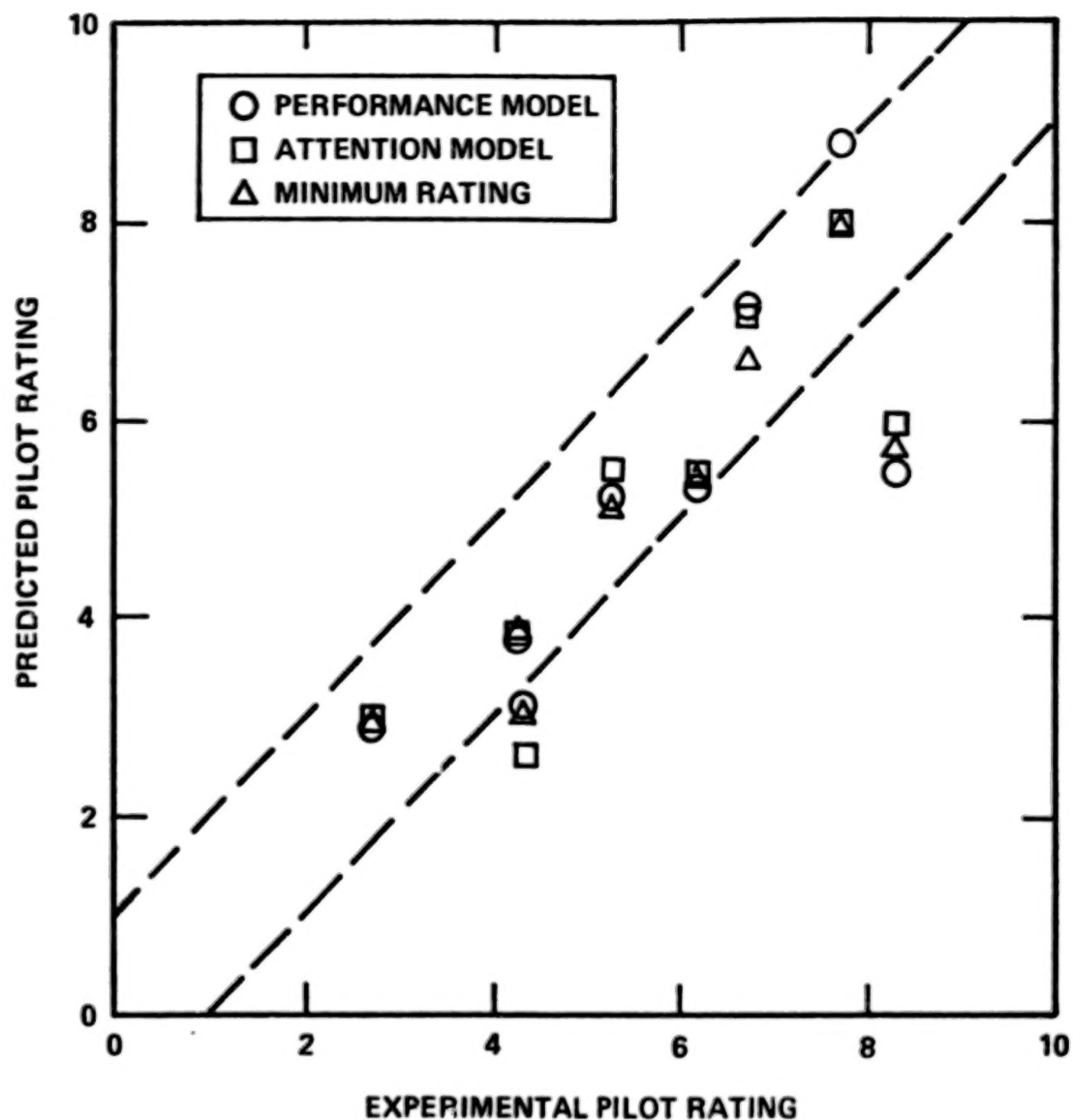
51

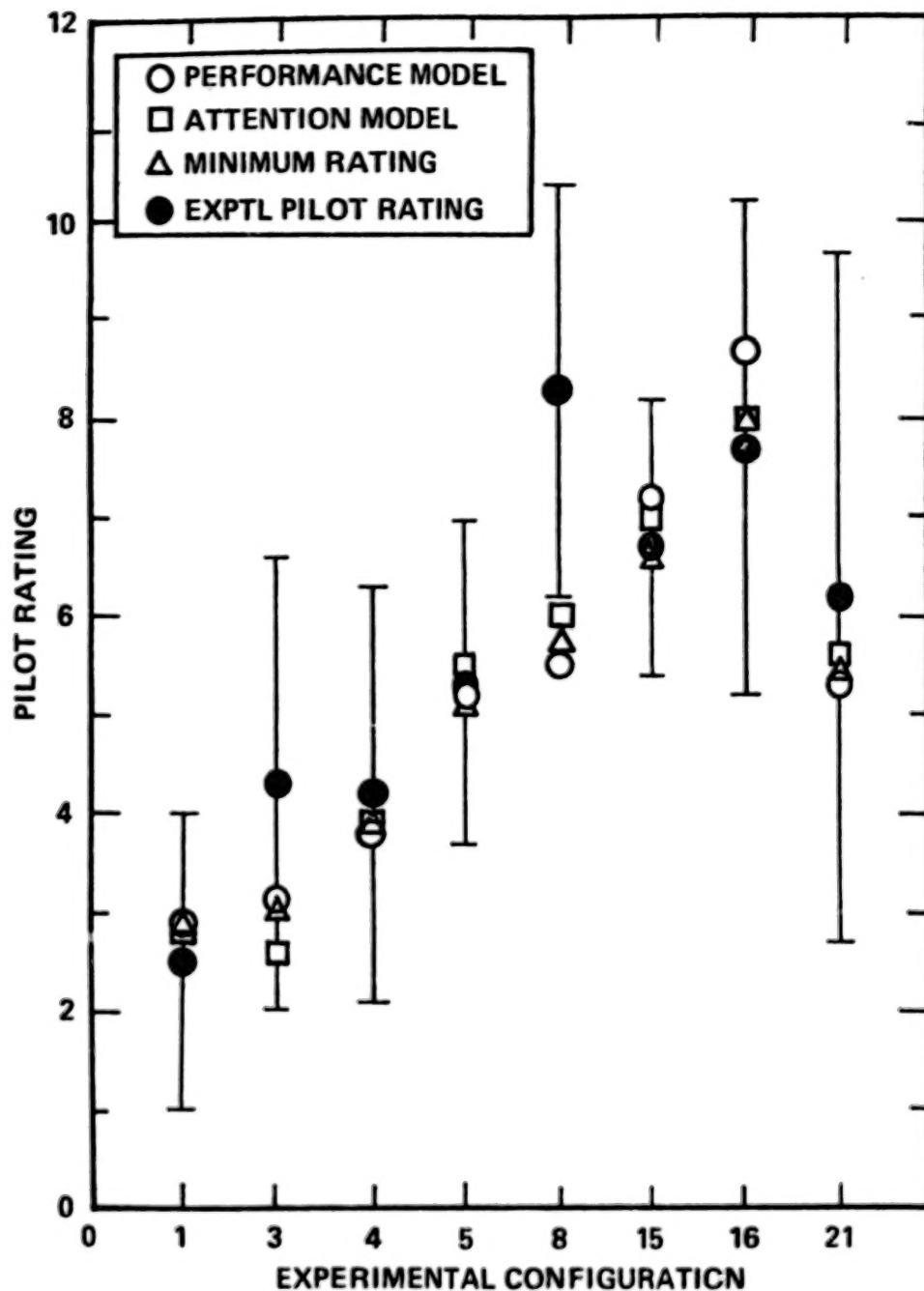Figure 6. - Predicted versus Experimental Pilot Ratings.

Figure 7. - Comparison of Predicted and Experimental Ratings.

Table 9

Definition of Handling Qualities Levels

Pilot Rating	H. Q. Level
1 - 3	1
3.01 - 3.99	1.5
4 - 6	2
6.01 - 6.99	2.5
7 - 9	3
9.01 - 9.49	3.5
9.5 - 10	4

Table 10

Comparison of Handling Qualities Criteria

Criterion\Config.	Handling Qualities Level								Error
	1	3	4	5	8	15	16	21	
Experimental	1	2	2	2	3	2.5	3	2.5	-
Performance Model	1	1.5	1.5	2	2	3	3	2	6
Attention Model	1	1	1.5	2	2	3	3	2	7
Minimum Rating	1	1	1.5	2	2	2.5	3	2	6
Bandwidth Model	1	2	2	2.5	1.5	2.5	3	2	5
$d\gamma/dV$	1	1	1	1	4	1	1	4	18
ω_{sp} vs n/α	1	4	4	4	1	4	4	2	22
ξ_{sp}	1	1	1	3	1	1	2	1	18
Combination	1	1.5	2	2	2.5	2.5	3	3	3

Compared in Table 10 are the three model-based rating expressions proposed in this study, the "bandwidth model" based on the work of Neal and Smith [2], three vehicle-centered criteria, and a criterion proposed by Rickard (the "combination" criterion) that combines the bandwidth model and the dY/dV criterion. All model-based criteria shown in this table provide substantially lower matching errors in terms of half-levels than the vehicle-centered criteria. Best results are obtained with the combination criterion; the bandwidth model and the three expressions based on the optimal-control model also perform well.

4.6 Discussion of Results

The generally good match between "predicted" and experimental pilot opinion ratings suggests that the model-based approach described in this report is basically valid. The technique is shown to replicate experimental results reasonably well across a set of conditions that spans a range of handling qualities levels and problems.

The schemes tested in this study replicate the experimental data nearly as well as other model-based schemes that have been tested against the same data base, and substantially better than the vehicle-centered criteria that have been tested. Because the procedure is based on a pilot/vehicle model of considerable generality and demonstrated validity, this scheme ought to be valid for other aircraft configurations and, with appropriate definitions of performance requirements, other flight tasks as well. Further study is required to compare the BBN techniques against other model-based procedures and to further compare the usefulness of the three rating expressions tested in this study.

Resources did not permit a detailed study of the inability to obtain a good match to the experimental rating for Configuration 8. The differences between the average ratings for Configurations 8 and 21 (which our prediction scheme predicts to be negligible) were apparently not due to training effects; these two

configurations were presented to the test pilots in a balanced order.*

It should be noted that all tests of the proposed methodology have been based on steady-state analysis appropriate to conditions at a single altitude. Although steady-state-like tasks were important determinants of pilot opinion, transient-response behavior was also important. There may have been some aspects of glide-slope capture and other transient maneuvers that were especially adverse for Configuration 8. Additionally, it is possible that a different choice of steady-state parameters (e.g., turbulence appropriate to a lower altitude, different "limits" on throttle response) may have differentiated between Configuration 8 and 21.

Data obtained in the Douglas study were used in the BBN study because of their applicability to large transports. As this study was undertaken well before the BBN analytical study was conceived, the data base is not sufficient for a thorough test of the model-based approach. Specific methodological deficiencies include:

1. Sparcity of Performance Measurements. Pilot opinion ratings were the only data published relating to closed-loop pilot/vehicle performance. Objective performance measures such as rms errors, pilot describing functions, spectra, or time histories are not available. Thus, we cannot determine the pilot's "operating point" in terms of pilot-related model parameters, and we cannot verify the ability of the model to predict objective performance.

2. Large Rating Variability. Standard deviations for pilot ratings, as determined across subjects, were relatively large, reaching a maximum of 3.5. Clearly, large variability in the data hinders a rigorous test of the prediction scheme.

---

* W. W. Rickard, personal communication.

To some extent, the large standard deviations resulted from a small subject population (only 3 subjects provided rat_igs for four of the configurations explored in the BBN study). As described below, a more significant factor may have been an insufficiently specific evaluation procedure.

3. <u>Insufficiently Specific Evaluation Procedure</u>. Typically, each pilot was allowed two "flights" per configuration: an initial flight without turbulence, and a follow-up flight with moderate turbulence. The pilots were encouraged to perform maneuvers that would aid in developing their rating, and they were asked for a single overall rating of the configuration at the end of the two flights. While all subjects appeared to consider the same basic maneuvers and subtasks (glide-slope capture, glide-slope tracking, recover from mistrim, open-loop vehicle response), we do not know the extent to which each pilot weighted the various response categories. Different weightings might have led to different ratings for the same configuration -- a possible explanation for the large pilot-to-pilot variability observed in this study. Differences in the pilot's expectations of system performance are an additional potential source of rating variability.

Consideration of these methodological shortcomings suggests alternative approaches in future studies, as outlined in the final .section of this report.

5. CONCLUSIONS AND RECOMMENDATIONS

A technique based on the optimal-control model for pilot/vehicle systems has been developed for predicting pilot opinion ratings. Three variations of this technique provide a good match to opinion ratings obtained in a manned simulation study of large commercial transports in landing approach.

The model-based technique developed in this study has a number of features which should enhance its applicability to other aircraft configurations and other flight tasks and should allow wider application than alternative handling qualities prediction schemes:

1. One is able to proceed in a straightforward manner from a description of the task environment and of task requirements to a prediction of pilot opinion ratings. The general form of the rating expression and of the underlying pilot model is invariant across applications.

2. No constraints are placed on the nature of the vehicle response, and the pilot model is relatively free form. Thus, "unconventional" aircraft dynamics may be considered. In order to minimize computational requirements, however, reduced-order modelling is recommended where predictions are not materially affected.

3. A scalar metric for attentional workload is expressed in terms of a model parameter related to the signal/noise properties of the pilot's response. Thus, the treatment of workload is independent of the details of the flight task.

4. The effects of display parameters, turbulence, and other environmental factors on pilot opinion rating are readily considered.

Encouraging results obtained with the model-based technique tested in this study warrant further research to provide a more rigorous test of the procedure and to determine its range of

validity. Such a study should be subjected to the following guidelines:

1. **Flight Test Standardization**. The flight tests performed for the purpose of obtaining pilot opinion ratings should be standardized so that all pilots perform the same maneuvers on the aircraft. Either separate ratings should be obtained for individual maneuvers, or care should be taken to assure that all pilots weight the various maneuvers in the same manner when assigning an overall rating to the aircraft.

2. **Define Performance Criteria**. Through a carefully prepared and administered questionnaire, subjective performance criteria should be determined for the various test maneuvers. If practical, test pilots should be encouraged to adopt a common set of criteria to minimize rating variability.

3. **Performance Measurement**. Objective measures of system performance and pilot response behavior should be obtained in addition to pilot opinion ratings to provide a more rigorous test of the method.

APPENDIX A

THE OPTIMAL-CONTROL PILOT/VEHICLE MODEL

A.1 Model Structure

The model is based on the assumption that the well-motivated, well-trained human operator behaves in a near optimal manner subject to his inherent constraints and limitations. A block diagram of the pilot-vehicle model is given in Figure A-1. The portion of the model which pertains specifically to the pilot is shown within the dashed line. Principal model elements are:

a. A linearized description of the vehicle dynamics given by the following state equation:

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{u}(t) + \underline{E}\underline{w}(t) \qquad (A-1)$$

where $\underline{x}(t)$ is the vector that describes the state of the vehicle, $\underline{u}(t)$ the pilot's (vector) control output, and $\underline{w}(t)$ a vector of independent white driving noise processes. (If the external forcing functions are rational noise spectra of first order or higher, the resulting "input states" are incorporated into the state vector $\underline{x}(t)$.)

b. An "output" (or display) vector, which, in general, consists of a linear transformation of the state variables and is given as

$$\underline{y}(t) = \underline{C}\underline{x}(t) \qquad (A-2)$$

c. A representation of the pilot's limitations by means of an equivalent perceptual time delay τ, an equivalent white observation noise vector $\underline{v}_y(t)$, and a white motor noise vector $\underline{v}_m(t)$. These two vector noise processes account for
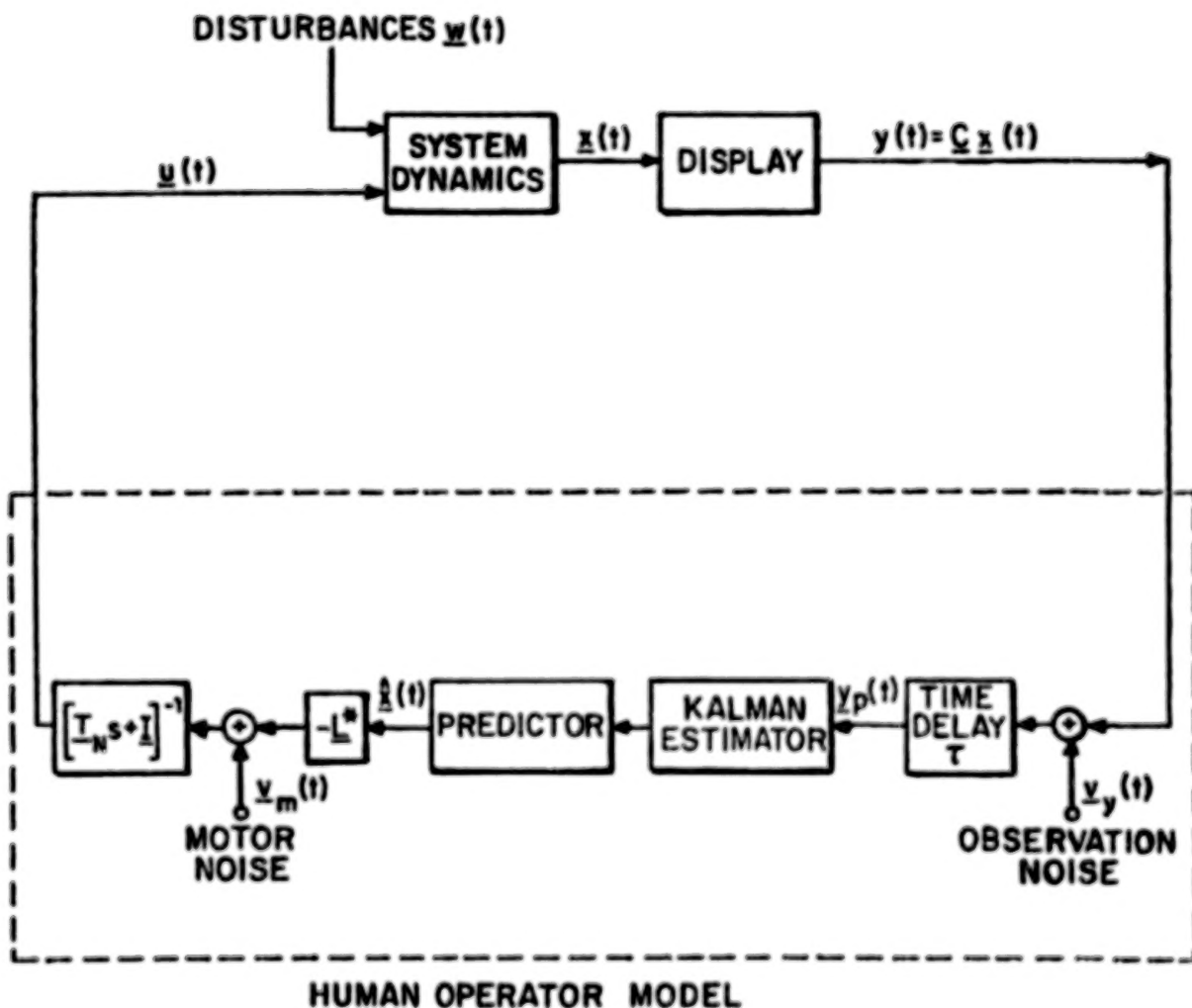
Figure A-1. - Optimal-Control Model for Pilot/Vehicle Systems.

61

pilot remnant.[*]
d. A least mean-squared predictor to compensate partially for the inherent time delay, and a Kalman filter to obtain the best estimate of the vehicle state.

e. A set of optimal gains $[-L^*]$ acting on the best estimate of the state vector, the output of which is the commandec control signal $\underline{u}_c(t)$.

The primary output variable of the optimal-control model is the pilot's control signal, $\underline{u}(t)$. Once the characteristics of this signal have been determined, prediction of various system and pilot performance metrics is relatively straightforward. For steady-state control tasks, predictions include variance scores for important system variables (e.g., mean-squared path, attitude, and control deviations) along with pilot describing functions and "remnant" spectra [10,11]. For tasks with deterministic (i.e., transient) inputs such as wind shears, ensemble statistics for time histories may be predicted [18].

The problem may be analyzed in a piecewise-steady-state manner when one or more statistical properties of the task changes slowly with time. Measures appropriate to steady-state tracking are predicted, but these predictions are considered as function of time, range, or altitude. An example of such a task is ILS tracking on final approach in the presence of zero-mean random turbulence, where typically the intensity and bandwidth of the gusts, the effective display gain associated with the ILS instrument, and the control objectives change with altitude and range.

The optimal predictor, optimal estimator, and optimal gain matrix represent the set of "adjustable parameters" by which the

---

[*] Observation noise appears to account for most of the measurable remnant. Motor noise is included in the model mainly to reflect the pilot's imperfect knowledge of the vehicle response behavior.

pilot tries to optimize his behavior. The general expressions for these model elements depend on system dynamics, according to well-defined mathematical rules that are described in [10,11]. For a given level of attention (reflected in the observation noise vector as described below), the controller is assumed to adopt a response strategy to minimize the following "cost function" (or performance index):

$$J = E \left\{ \lim_{T \to \infty} \frac{1}{T} \int_0^T \left[\sum_{i=1}^{N_y} q_i \, y_i^2(t) + \sum_{i=1}^{N_u} (r_i \, u^2(t) + g_i \, \dot{u}^2(t) \right] dt \right. \quad (A-3)$$

A.2   Display Analysis

A basic assumption of the optimal control model of the human operator is that the human perceives a noisy, delayed version of the displayed variables; i.e., if $\underline{y}_y(t)$ is the vector of perceived variables

$$\underline{Y}_p(t) = \underline{Y}(t-\tau) + \underline{y}_y(t-\tau) \quad (A-4)$$

where $\underline{y}$ is defined by (A-2). The human's time delay τ is a parameter of the model. Comparison of model results with experimental data for a variety of systems and input conditions has yielded values of $\tau = 0.2 \pm 0.05$ seconds [13-15], numbers that are consistent with human time delays reported by others [27].

The observation noise vector $\underline{v}_y$ is an important part of the model. It is, essentially, a lumped representation of human randomness. On the basis of considerable experimentation, a relatively simple set of rules for specifying $\underline{v}_y$ have been determined [14]. Each component of $\underline{v}_y$ is assumed to be a zero-mean, white (i.e., wide-band with respect to the variables on the displays), Gaussian noise process that is uncorrelated with other such noise processes and with system input noise

63

disturbances. Therefore, each noise process can be quantified by a single parameter, namely, its autocovariance. For manual control situations in which the displayed signal is large enough to negate the effects of human resolution ("threshold") limitations, the autocovariance of each observation noise component appears to vary proportionally with mean squared signal level. Thus,

$$E\{v_{yi}(t)v_{yi}(s)\} = v_i(t)\delta(t - s)$$

$$V_i(t) = \pi P_i E\{y_i^2(t)\}$$

(A-5)

where P_i is the "noise/signal ratio" and has units of normalized power per radian per second. Numerical values for P_i of 0.01 (i.e., -20 dB) have been found to be typical of a variety single-variable control situations [10,13]. The relative invariance of P_i with control task parameters suggests that the basic observation noise defined by (A-5) represents a processing limitation of the human pilot.

When display characteristics are not ideal it is necessary to modify the expression for the observation noise covariance associated with a particular display variable. In the work thus far, prime attention has been devoted to two display limitations: namely, threshold limitations and the lack of a zero reference. These phenomena are accounted for by letting the autocovariance for each observation noise process be

$$V_i(t) = P_i \left[\frac{\sigma_i^2}{K_i^2(\sigma_i,a_i)} + \sigma_{io}^2 \right]$$

(A-6)

where the subscript i refers to the ith display variable. The quantity $K(\sigma_i,a_i)$ is the describing function gain associated with a threshold device

64

$$K(\sigma,a) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{-(a/\sigma\sqrt{2})} e^{-x^2} dx \qquad (A-7)$$

where a is the threshold and σ is the standard deviation of the "input" to the threshold device.* This threshold gain causes the observation noise covariance to become greater as the signal becomes smaller relative to the threshold. The expression in (A-7) assumes that a zero reference for y_i is available. Otherwise, the observation noise is assumed to scale with the mean squared deviation of $y_i(t)$ from its nearest reference indicator. This factor is accounted for by the "residual-noise" covariance σ_{io}^2 in (A-6). If the desired deviation from the reference is d and it is assumed that $E_{\{y_i(t)\}} = a$, then $\sigma_{io} = d$.

A.3 Model for Attention-Sharing

The model for attention-sharing described in [13,21] is reviewed here.

We consider attention-sharing to be required at three levels: between manual control and non-control tasks; between subtasks within the manual control task; and between displays associated with performing a given subtask. For example, a pilot might share attention between control and communication, between longitudinal and lateral control, and between flight path and attitude displays. Thus, define

f_t = fraction of attention devoted to the control task as a whole;

f_s = fraction of attention devoted to subtask s;

f_i = fraction of attention devoted to the ith display

---

* For nonzero mean signals this expression must be modified.

The effects of attention-sharing are incorporated in the following model for noise/signal ratio:

$$P_i = P_o \; \frac{1}{f_t} \; \frac{1}{f_s} \; \frac{1}{f_i} \tag{A-8}$$

where P_i is the noise/signal ratio associated with the ith display when attention is being shared and P_o is a baseline noise/signal ratio associated with a high-workload single tracking task.

Typically a value of -20 dB is associated with P_o, as it is representative of performance in a variety of single-variable tracking tasks. This is not an absolute limit on human performance, however; control tasks especially sensitive to human operator randomness have yielded substantially lower noise/signal ratios [13,15]. Thus, attention is defined relative to that which is typically achieved in a "standard" laboratory tracking task.

A.4 Independent Model Parameters

Once the equations of motion of the physical system have been quantified as in Eqs. (A-1) and (A-2), additional model parameters must be specified to reflect performance requirements (both objective and subjective), physical limitations of the system, and information-processing limitations of the human pilot. Most of these parameters fall into two categories: cost weightings, and display-related parameters.

A.4.1 Cost Weightings

The pilot is assumed to adopt a control strategy that minimizes a weighted sum of mean-squared (or integral-squared) response variables. In the case of landing approach, this performance index will generally include path, altitude, and control variables. Weightings are derived by first associating a maximum allowable value (or "limit") with each variable and then setting the corresponding cost coefficient equal to the square of the reciprocal of the limit. This scheme for selecting cost

66

coefficients has been used by Hess in his development of a handling qualities predictor [9] and by this author in previous analytical studies of landing approach performance [18,28].

A.4.2 Display-Related Parameters

As implied by the preceding sections on display analysis and attention-sharing, a number of model parameters must be specified to quantify the ability of the pilot to obtain relevant information from the flight-control displays. Three parameters must be determined for each perceptual variable utilized by the pilot: an effective perceptual "threshold", a "residual noise", and the fraction of attention allocated to that variable.

Two types of thresholds are considered: an "indifference threshold" that reflects subjective control requirements, and thresholds resulting from visual resolution limitations. In a landing approach task, indifference thresholds are typically associated only with outer-loop variables; for example, the pilot may decide not to attempt to reduce height error if his glide-slope indicator shows an error of less than 1/4 "dot". Similarly, he might not attempt to reduce speed error if it is within a single calibration mark of the reference speed.

In general, the pilot is assumed to obtain both displacement and rate information from a display indicator. For a symbolic display element that provides a cue proportional to its displacement from a reference, the resolution-related threshold may be computed as

$$a = \frac{D}{G} \frac{a_o}{57.3}$$

$$(A-9)$$

where "a" is the threshold in problem units, a_o is the perceptual threshold in visual units, D is the viewing distance, and G is the display gain relating display deflection (same units as D) to "error" (problem units). Non-symbolic and discrete displays would require a modified treatment.

67

APPENDIX B

QUESTIONNAIRE ADMINISTERED TO DOUGLAS AIRCRAFT TEST PILOTS

Name_____   Date_____   Page 1

QUESTIONNAIRE FOR DAC TEST PILOTS

Introduction

Douglas Aircraft Company and Bolt Beranek and Newman of Cambridge, Mass. are currently working on a joint project for NASA-LRC to develop techniques for predicting aircraft handling qualities. The data obtained in the DAC experiments of 1975 on longitudinal flying qualities will be used to develop and test these techniques. Because the proposed prediction schemes will be based on a quantitative pilot/vehicle model, it is important to obtain a precise and quantitative definition of the flight task, particularly with regard to the pilot's expectations in performing the landing approach. Your co-operation in answering this questionnaire -- designed to elicit this information -- will be appreciated.

It is assumed that your ratings were based on the so-called "Cooper-Harper" rating scale, or a scale similar to that one. Ratings 1-3 correspond to your ability to achieve "desired" performance with negligible to minimal pilot compensation (i.e., workload). Ratings 4-6 reflect situations ranging from desired performance with moderate compensation to adequate performance with extensive compensation. Ratings 7-9 correspond to situations where control is maintained but adequate performance is not achieved, even with considerable compensation. A rating of 10 reflects loss of control.

Question 1

It is assumed that lateral-directional handling qualities were quite satisfactory and that your pilot ratings were influenced primarily by longitudinal handling characteristics.

Is this assumption valid?   Yes_____   No _____

Question 2

The first phase of the approach was flown by panel instruments, the second phase via a simulated real-world display. Which phase was, in general, more important in determining your rating?

Phase 1 (instruments)_____   Phase 2 (real-world)_____

Question 3

Please identify aspects of the approach task (e.g., glide-slope capture, regulating against turbulence) that were important factors in determining ycur ratings. Also identify any self-imposed tasks that you would generally perform in the process of evaluating longitudinal-axis capabilities during final approach (e.g., rapid acquisition of a new pitch attitude, intentional deviation from and recapture of the beam). List the more important tasks first.

(1)

(2)

(3)

(4)

(5)

(6)

Question 4

The objective of this question is to determine your expectations of pilot/vehicle performance as quantitatively as possible for the subtasks defined above in Question 3. On the following pages, please list the "desirable" and "adequate" levels of response for variables that you consider to be important. When applicable, and wherever possible, try to specify values for variables relating to height error, sinkrate or path angle, attitude (displacement and/or rate), airspeed error, and control (displacement, force, and/or rates of change of these variables).

By "adequate" performance is meant the performance level that is barely acceptable (i.e., the boundary between a pilot rating of 6 and a rating of 7), whereas "desired" performance corresponds to a rating of 1.

This question is intended strictly to determine the performance aspects of what determines a pilot rating. The workload aspects are hopefully obtainable from the commentary that was obtained at the time of the evaluation experiments.

The variables you choose to quantify, and the way in which you choose to describe "desirable" and "adequate" performance levels, will depend on the particular subtask. For example, if you are considering glide-slope tracking, you might specify adequate height regulation as maintaining the aircraft within plus or minus 50 feet of the desired glide path 90% of the time. "Desired" performance might be maintaining the same bounds 99% of the time, or maintaining some tighter bound 90% of the time. Alternatively, you might specify the maintenance of a certain rms deviation about the desired glide path. (Note: these numbers are given only by way of exemplifying the method and are not intended to influence your specification of desired and adequate performance levels.)

Performance specifications for a transient maneuver would be expressed differently. For example, if you are considering the task of acquiring a new pitch attitude, you might specify adequate performance as accomplishing 90% of the maneuver in "t" seconds, and/or accomplishing the maneuver with less than "x" percent pitch overshoot using normal piloting procedure.

Be sure to specify the units for each variable (e.g., whether height error is in terms of feet, angular deviation from the glide path, or dots on the glideslope display). Please use a separate answer sheet for each subtask (defined in Question 3) for which you can define desired and adequate performance levels.

Name_____ Date_____ Page___

Subtask_____

Give the units, "desired" performance level, and "adequate" performance level for each important response variable. Use additional answer sheets for this subtask if necessary.

REFERENCES

1. "Flying Qualities of Piloted Airplanes". Military Specification, MIL-F-8785B (ASG), 1969.

2. Neal, T. P. and R. E. Smith: "An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes", Wright-Patterson Air Force Base, Ohio, Volume I, AFFDL-TR-70-74, 1970.

3. Onstott, E. D. and E. P. Salmon: "Airplane Flying Characteristics in Turbulence". Wright Patterson Air Force Base, Ohio, AFFDL-TR-70-143, 1971.

4. Black, G. T.; D. J. Moorhouse; and R. J. Woodcock: "Proceedings of AFFDL Flying Qualities Symposium Held At Wright State University", AFFDL-TR-78-171, December 1978.

5. Anderson, R. O.: "A New Approach to the Specification and Evaluation of Flying Qualities". Wright-Patterson Air Force Base, AFFDL-TR-69-120, 1970.

6. Arnold, J. D.; R. B. Johnson; and J. D. Dillow: "Pitch Paper Pilot Revisited". Proceedings of the Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1973.

7. Naylor, F. R.; J. D. Dillow; and R. A. Hannen: "Roll Paper Pilot". Proceedings of the Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1973.

8. Stone, J. R.: "Paper Pilot Ponders Supersonic Transports", Proceedings of the Ninth Annual Conference on Manual Control. Massachusetts Institute of Technology, Cambridge, Massachusetts, 1973.

9. Hess, R. A.: "Prediction of Pilot Opinion Ratings Using an Optimal Pilot Model". HUMAN FACTORS, 1977, 19(5), 459-475.

10. Kleinman, D. L.; S. Baron and W. H. Levison: "An Optimal Control Model of Human Response, Part 1: Theory and Validation". Automatica, Vol. 6, 1970.

11. Kleinman, D. L.; S. Baron and W. H. Levison: "A Control Theoretic Approach to Manned-Vehicle Systems Analysis". IEEE Trans. on Auto. Control, Vol. AC-16, No. 6, Dec. 1971.

12. Schmidt, D. K.: "Optimal Flight Control Synthesis via Pilot Modeling". Journal of Guidance and Control, Vol. 2, No. 4, Jul.-Aug. 1979.

13. Levison, W. H.; J. I. Elkind and J. L. Ward: "Studies of Multi-Variable Manual Control Systems: A Model for Task Interference". NASA CR-1746, May 1971.

14. Levison, W. H.: "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant". Wright-Patterson Air Force Base, Ohio, AMRL-TR-70-93, March 1971.

15. Zacharias, G. L. and W. H. Levison: "A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies". AMRL-TR-79-17, March 1979.

16. Kleinman, D. L. and S. Baron: "Analytic Evaluation of Display Requirements for Approach to Landing". NASA CR-1952, Nov. 1971.

17. Kleinman, D. L. and W. R. Killingsworth: "A Predictive Pilot Model for STOL Aircraft Landing". NASA CR-2374, March 1974.

18. Levison, W. H.: "Analysis and In-Simulator Evaluation of Display and Control Concepts for a Terminal Configured Vehicle in Final Approach in a Windshear Environment". NASA Contractor Report 3034, August 1978

19. Levison, W. H.; S. Baron; and A. M. Junker; "Modeling the Effects of Environmental Factors on Human Control and Information Processing". AMRL-TR-76-74, August 1976.

20. Lancraft, R. E.; and D. L. Kleinman: "On the Identification of Parameters in the Optimal Control Model". Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, Ohio, March 20-22, 1979.

21. Levison, W. H.: "A Model for Mental Workload in Tasks Requiring Continuous Information Processing". Mental Workload Its Theory and Measurement, Plenum Press, New York and London, 1979.

22. Wewerinke, P. H.: "Human Operator Workload for Various Control Situations". Tenth Annual Conference on Manual Control, Wright-Patterson Air Force Base, Ohio, 1974.

23. Volkmann, F. C.; A. M. L. Schick; and L. A. Riggs: "Time Course of Visual Inhibition During Voluntary Saccades". Jl. Opt. Soc. Am., 58, pp. 562-569, 1968.

24. Cooper, G. E.; and Harper, R. P., Jr.: "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities". NASA TN D-5153, April 1969

25. Rickard, W. W.: "Longitudinal Flying Qualities in the Landing Approach". Proceedings from the Twelfth annual Conference on Manual Control. NASA TMX-73,170, May 1976.

26. Levison, W. H.; S. Baron: "Analytic and Experimental Evaluation of Display and Control Concepts for a Terminal Configured Vehicle". Report No. 3270, Bolt Beranek and Newman Inc., Camb., MA July 1976.

27. McRuer, D. T.; D. Graham; E. S. Krendel; and W. Reisner, Jr.: "Human Pilot Dynamics in Compensatory Systems: Theory, Models and Experiments with Controlled Element and Forcing Functions Variations." AFFDL-TR-65-15, July 1965.

28. Baron, S. and W. H. Levison: "A Manned Control Theory Analysis of Vertical Situation Displays for STOL Aircraft". NASA CR-114620, April 1973.

1. Report No. NASA CR-3257	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A Model-Based Technique for Predicting Pilot Opinion Rating for Large Commercial Transports		5. Report Date April 1980
		6. Performing Organization Code
7. Author(s) William H. Levison		8. Performing Organization Report No.
		10. Work Unit No.
9. Performing Organization Name and Address Bolt Beranek and Newman Inc. Cambridge, MA 02138		11. Contract or Grant No. NAS1-15529
		13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Army Project No.

15. Supplementary Notes

Langley Technical Monitor: Martin T. Moul
Final Report

16. Abstract

A model-based technique for predicting aircraft handling qualities is described. Features of this procedure, which is based on the optimal-control model for pilot/vehicle systems, include (1) capability to treat "unconventional" aircraft dynamics, (2) a relatively free-form pilot model, (3) a simple scalar metric for attentional workload, and (4) a straightforward manner of proceeding from descriptions of the flight task environment and requirements to a prediction of pilot opinion rating. The method was able to provide a good match to a set of pilot opinion ratings obtained in a manned simulation study of large commercial aircraft in landing approach.

17. Key Words (Suggested by Author(s)) Flying Qualities Pilot Modeling Flight Simulation Stability and Control	18. Distribution Statement Unclassified - Unlimited Subject Category 08
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 75	22. Price* $5.25

END

June 30, 1981